Home-Start Early Speech and Language Study

Phase 1 evaluation report

James Law

Jenna Charlton

Newcastle University

Robert Rush

Queen Margaret University Edinburgh

JULY 2018

Funded by Department for Education

nesta

Department for Digital, Culture Media & Sport

Foreword

The vital importance of early speech and language acquisition is now becoming widely recognized. Damian Hinds, Secretary of State for Education, while talking about social mobility recently said "*Children with poor vocabulary at age five are more than twice as likely to be unemployed when they are aged 34.It's command of language, being able to express ourselves effectively, that is the gateway to success in school – and later on into later life*" (July 2018).

When we think of child malnourishment we think of children going hungry or having a poor diet because we know that children need food to develop. Now we are also thinking in terms of young children needing to hear and use words and sounds to help their brains develop. Children hearing less than 21,000 words a day are not getting their optimum diet – they are also malnourished.

This report provides encouraging preliminary evaluation findings from an innovative study using technology to give parents individualised feedback on their interactions with their children. It reports on the first pilot study of home based volunteers trained to coach parents to expand talk/interactions with their children. The study shows that the intervention does appear to make a difference to parental behaviour. While only working with small numbers of families over a short period of time, the evaluation demonstrated statistically significant positive outcomes from comparing parents before and after the intervention for adult word count, the Developmental Snapshot and the subscales of The Home Inventory were achieved. The sample size was too small to detect effects on children, future evaluations are planned to assess this.

Most importantly parents liked the project. They valued the personalised data reports specifically about their family and the customised support around what they needed to change. They could see the impact on their children and their wider family life. These are key building blocks to have tested and have in place to ensure that parents are able to encourage their children to acquire a wider range of vocabulary early in their lives.

The families in this project have complex lives and were managing some real challenges from living in poverty, being asylum seekers, having children on child protection plans and living with mental health issues. To demonstrate positive findings for families in these circumstances is particularly pleasing as the project aimed to target families in areas of deprivation and where local authority school attainment figures were poor. It demonstrates social mobility in action. While very important to be able to reach and support families who are living complex lives it is not necessarily the case that the language environment in the home is not rich. It is also not always the case that those who are considered to be more affluent have a richer language environment. Going forward our aim is to support a range of families to improve their interactions, each starting from their own individual baseline.

There are some key lessons to be learnt from this pilot and as the programme is retested across the UK these factors can be incorporated into its implementation. What is clear from this report is that adult talk must be in the form of conversational turns to increase child vocalisations. Indeed, Gilkerson et al (2018) have recently published the results of a longitudinal study using LENA showing conversational turns at 18-24 months predicted scores in language, comprehension and expression tests at ages 9-13 years. The training for the volunteer coaches has been updated to ensure that there is a clear focus on the importance of turn taking and 'scaffolding' in conversation and emphasis will be placed on the quality of the interaction to improve child outcomes .

We are committed to working with the local Home-Starts, the volunteer coaches and the families to understand the parent perspective in more detail and to identify what is causing the 'noise' identified in this report from within group variation in findings over the programme. The LENA Foundation have during the course of this pilot recommended a revised approach to LENA HOME which involves fewer recordings, and a greater focus on the activities the home-visitor carries out with the family. LENA recognised that more guidance on the programme content led to better fidelity and that the coaching sessions needed consistency to have the greatest impact. It is anticipated that as this is adopted less within family variation will be identified.

In summary this first stage evaluation demonstrated that the quality and quantity of stimulation and support available to the child increased, the parents' perception of their child's development improved and the amount the parents spoke to their children increased. This is a great platform from which to further develop this project.

GLOSSARY

Adult word count; LENA data, frequency of adult words spoken

Automatic vocalisation assessment; LENA data about the expressive language development of a child

Between Group; Refers to comparisons of data between the Home Start 'intervention' group, and the control group

Child Vocalisation Count; LENA data, frequency of child words, babbles, and pre-speech communicative sounds

Conversational turns; LENA data, frequency of vocalisation turns between an adult and child

Correlation; Type of analysis exploring the relationship between two or more variables

Developmental Snapshot; LENA assessment, a 52-item parent-completed evaluation of language skills for infants and toddlers focusing on well-established milestones associated with expressive and receptive language skills

Digital Language Processor (DLP); LENA equipment, a small digital data recorder that records the child's language environment

HOME Inventory; HOME is an acronym for Home Observation for Measurement of the Environment

LENA (Language Environment Analysis); LENA Pro software, automatically analyses and segments the audio data, providing four primary reports plus a composite report that can be viewed in monthly, daily, hourly, and five-minute time-frames

LENA HOME; Intervention programme involves trained coaches supporting parents to expand the frequency and range of conversations held with their young children

Mixed ANOVA; Type of analysis, analyses change in data over time in one group, and compares data between two groups (Home Start 'intervention' group and the control group)

Preschool Language Scale (PLS-UK); A comprehensive developmental language assessment that measures a child's expressive and receptive language ability

Regression; Type of analysis, explores the predictive relationship of variables on an outcome variable

Standardised; Term used when an assessment has population-level data that provides 'normed' data for comparison to sample data (e.g. mean scores, percentiles)

Within Group; Refers to data of one group of participants, the Home Start 'intervention' group

EXECUTIVE SUMMARY

Aims

This report summarises the preliminary findings of an intervention feasibility study carried out by Home-Start UK involving families from four local Home-Starts in England. The aim of the study was to use the LENA technology, the so called "word pedometer", language measures and measures of the home environment over a twelve-week period to understand the impact of weekly home visits by a trained Home Start volunteer on children's home environment, children's language ability and parent-child verbal interactions.

The study

The participants

35 families were recruited by four local Home-Starts in England for the Lena HOME programme. This report focuses on 21 families for whom there was a full set of 12 recordings and 21 "controls" from the LENA database in the US matched for age, gender, and maternal education (banded so that the UK and the US levels correspond).

The measures

The evaluation collected data from the Preschool Language Scale-5 about children's language ability, data from the HOME Inventory about the quality and quantity of stimulation and support available to a child in the home environment, and LENA data which measures the child's language environment (number of vocalisations from the child and an adult, number of conversational turns, proportion of electronic noise as well as a 'developmental snapshot', a measure of a child's expressive and receptive language development). Comparison 'control' LENA data was also collected from the US-based LENA database.

The findings

Over the three months a number of the HOME Inventory scores changed in a positive direction indicating that the home environment was sensitive to the intervention. A range of different factors predicted these home scores of which the most consistent of those we measured was maternal education. The Preschool Language Scale expressive and receptive language raw scores changed positively but the data were very variable - "noisy" – the range of the scores was wide and the changes to the standardised scores – i.e. those scores taking into account the children's age (standardised scores and percentiles ranks) did not change significantly. The LENA~ Developmental Snapshot changed positively on all scales but only the developmental age and the count went up significantly. Standardised and percentile rank scores went up but not enough to affect statistically significant change. Of the LENA measures taken from the LENA body worn device the scale which seemed to respond most to the intervention was the Adult Word Count [AWC] In other words, parents were talking more. Conversational turns and Automatic Vocalisation Assessment changed in the right direction but not significantly. We did find that if we compared the baseline score with the average of week 2-12 that there was a significantly different change, again for AWC. Yet it is important to note that there was a considerable variability from week to week and not necessarily in an upward direction. Parents were asked to comment on their experiences and we include some quotes from positive responders plus two short case studies which confirm the positive experience of the intervention as far as the parents in question were concerned. From this we conclude that the LENA focus on parents does appear to be making a difference to the HOME Inventory scores, to the parents' perception of their child's development and to the amount that the parent speaks. Within these data, although the data suggest that changes are going in the right direction, this has not translated into significant changes in the children's language test scores or the child interaction behaviours.

The comparisons with the US data suggested that the UK children had, overall, rather higher scores that the US children (with a moderate effect size of approximately 0.4). Perhaps unsurprisingly the LENA variables were associated with one another and this picture was stronger in the UK intervention group relative to the US sample This might be anticipated given that this might have been expected to be an outcome of the intervention – i.e. to increase the relationship between the amount that the adult speaks and the number of turns that are taken. Beyond this there was no significant difference on any of the three key LENA variables between the groups.

The results of each analyses are summarised in the table below; positive outcomes are represented by P, statistically significant outcomes are represented by S.

Analyses and Variables	**Within group analyses**	**Between group analyses**	**Analysis #**
HOME			
Home Data; Pre- and post-intervention associations	S		1
Home Subscale Change:			
Responsivity	S		2
Acceptance	S		3
Organisation	S		4
Learning Materials	S		5
Involvement	P		6
Variety	S		7
Total	S		8
HOME Regression	S		9
LANGUAGE			
Language Test Data; Pre- and post-intervention scores	S		10
Language Test Data; Difference between time points:			
Comprehension raw score	S		11
Comprehension standard score	P		12
Comprehension percentile score	P		13
Expression raw score	S		14
Expression standard score	P		15
Expression percentile score	P		16
Language Total raw score	P		17
Language Total standard score	P		18
Language Total percentile score	P		19
Language Test Regression	P		20
LENA			
Developmental Snapshot (Age)	S		21
Boxplots	P		22
Descriptive Statistics	P		23
Change Over Time; Adult Word Count (AWC)	S	P	24
Change Over Time; Conversational Turns (CT)	P	P	25
Change Over Time; Child Vocalisation Count (CV)	P	P	26
Change Over Time; Automatic Vocalisation Assessment (AVA)	P		27
Change Over Time; TV (Secs)	P		28
LENA Case Studies	P		29
The Comparison Group			
Adult Word Count (AWC)		P	30

Child Vocalisation Count (CV)		P	31
Conversational Turns (CT)		P	32
Correlations Between Variables	S	P	33
LENA Regression	P		34

Conclusions

The assessment data reported here suggest that it may be easier to detect changes on the Home environment and the adult word counts than it is in in the child specific behaviours. This probably reflects this LENA programme's emphasis on parental behaviour. The LENA data provide an incomparable level of detail about the parent/child interaction process at regular intervals but this detail, in this study at least, demonstrates clearly that there is a great deal going on in these patterns of interaction of which the LENA Home intervention is only one element. There is potential to develop this further, but we would suggest that we need to better understand the parental perspective on the intervention and about the contribution of the volunteers. It would be important to have more feedback from the volunteers about exactly what the messages were that they were giving in the home and how they felt the parents and carers responded to the feedback and advice. The biofeedback element of LENA provides both an unrivalled level of detailed data and a powerful tool for instruction at an individual level, but these results would suggest that we are some way off translating this individualised model of intervention delivered by volunteers into results showing that the intervention works at a group level.

INTRODUCTION

BACKGROUND

Social disadvantage is known to impact on children' language development in the preschool years (Hoff 2006; Huttenlocher *et al.* 2010; Law *et al.* 2011; Roy and Chiat 2013*)* and in later adolescence (Spencer *et al.* 2012). Hart and Risley's seminal work *Meaningful Differences in the Everyday Experiences of young American Children* (Hart and Risley 1995) has been instrumental in driving an understanding in the US about the widening gap in the language skills of very young children in white collar, blue collar and welfare families. This, in turn, has led to the Big Word Gap campaign http://www.bwgresnet.res.ku.edu/ and the Thirty Million Word Initiative (http://thirtymillionwords.org/) which have become so salient that they are now supported by the Clinton Foundation (Clinton Foundation 2013) (https://www.clintonfoundation.org/blog/2013/10/03/closing-word-gap). Although the concerns are undoubtedly real there is a danger of catastrophizing the issue based on a study which only included 42 children, albeit 42 children whose language had been very closely observed for the first three years of their lives. However, there are clearly a great many other issues about the nature of parental input relative to broader issues associated with material deprivation which we need to better understand if we are to move from a recognition of a concern to evidence based recommendations about what, if anything, should be done to support children's early language development. For example, one of the key questions for researchers, and indeed practitioners, is how much variance is explained by parental input relative to broader social risks such as income poverty, social class or social deprivation. In an important relatively recent study Huttenlocher *et al.* (2010) suggested that caregiver speech to children partially mediated the relationship between social background and the children's vocabulary development.

Young children's language experiences and language outcomes are highly variable (Schwab et al. 2016). Research in recent decades has focused on understanding the extent to which family socioeconomic status (SES) relates to parents' language input to their children and, subsequently, children's language learning. A common early intervention approach for preschool children with language problems is parent-child interaction therapy (PCIT) (Klatte and Roulstone 2016). PCIT has positive effects for children with expressive language problems. It appears that speech and language therapists (SLTs) conduct this therapy in many different ways. Parent-delivered home programmes are frequently used to remediate speech and language difficulties in young children (Tosh et al. 2017).

The outcomes for such interventions can prove a challenge of measure because, while it is perfectly possible to assess a child oral language skills, it can be difficulty to capture the to and fro of interaction between parent and child and more difficult to assess whether this interaction is changing in the predicted direction following intervention. One type of measurement of interaction which has attracted considerable attention recently grew out of the work of Hart and Risley (1995) are body worn audio recorders capable of capturing interaction in real time. A recent systematic review of the use of LENA (Language ENvironment Analysis) technology (Wang et al.2017), addressed three research questions: (a) What types of studies have been conducted, and with which populations, since the launch of LENA technology? (b) What challenges related to use of LENA technology were identified? (c) What are the implications for practice and future research using LENA technology? Electronic databases, the LENA Research Foundation website, and bibliographies of already-included studies were searched; 38 studies were identified. The authors selected studies on the basis of purpose, design, participant characteristics, application of LENA technology, and results. They found that LENA technology was used with a range of populations to yield a variety of information. Though challenges and limitations are associated with LENA technology, great potential exists for further research and a resultant increase in evidence-based understanding of early language development and interventions on its behalf.

It is important to acknowledge that changing parental behaviours in this way is rarely straightforward. The behaviours can be difficult to detect, children are likely to be very variable across different time periods, external events may be in play which affect performance, but over which researchers have no control or indeed may not be aware of. While many targeted interventions using, for example, milieu therapy have shown positive results (Robert and Kaiser 2011), this is not always the case. A trial, specifically of LENA, as an intervention in the US did show effects (Suskind et al 2016) but a recent trial of parenting interventions, albeit not using LENA, to promote language in Victoria, Australia did not show any effects (Wake et al.2017) and a recent trial in the UK which included LENA when used with very young children had an immediate effect but this was not sustained (McGillon et al. 2017). So the evaluation of intervention programmes designed to demonstrate the malleability of parental interactive behaviours remains an issue for careful research.

What is Home-Start?

Home-Start is one of the leading family support charities in the UK. Home-Start volunteers help families with young children deal with the challenges they face. Last year Home-Start supported 60,000 children in 30,000 families, in communities across the UK. Home-Start is committed to providing real evidence of the difference volunteers' support makes for families and they measure the issues faced by families, and the changes to their situation while they are working with a volunteer. Early language development is key for the children in these families and accordingly Home-Start is working with the American organisation, **the LENA Foundation**, with funding from the Department for Education, **NESTA** and the Department for Culture, Media and Sports.

The LENA project

What is LENA?

'LENA' stands for Language ENvironment Analysis. The LENA system includes a Digital Language Processor (DLP), a small digital data recorder worn by the child which records the language environment of a child throughout the day; sometimes referred to as a 'word or verbal pedometer'. Data is protected by allocating a unique ID to the child, and the DLP cannot play back recordings from the day or interpret any meaning of vocalisations/words. Data from the DLP is uploaded by USB port onto the LENA Pro software on a laptop or computer and processed automatically with no need for human intervention. LENA provides data on patterns of speech, or 'vocalisation count' data (number of child vocalisations, number of adult words, number of conversational turns between adult and child) and environment data (proportion of electronic noise) via an Audio Processing System which comprises four components: information flow, information processing, algorithmic processing models, and professional human transcriptions. Acoustic properties of the audio are segmented by algorithmic models to identify sounds of varying amplitude and intensity, and feature extraction identifies the source of the audio signal through iterative modelling in order to categorise the sounds into key LENA outcomes. Each recording results in a set of graphs that shows interactions by hour and across the day. The LENA Pro software provides four primary reports plus a composite report that can be viewed in monthly, daily, hourly, and five-minute time-frames. Export features and a sophisticated data mining tool allow you to customize analysis with ease, down to the millisecond. The LENA System™ is the industry standard for measuring talk with children from birth to three years, a critical factor in early brain development.

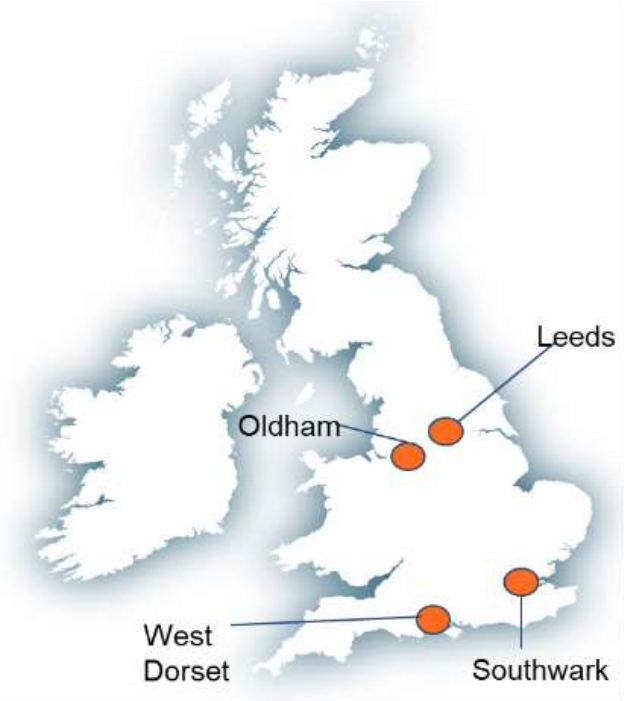
Lena Home in the UK

LENA HOME™ involves trained coaches supporting parents to expand the frequency and range of conversations held with their young children.

Families are given free access to a 'word pedometer' device, worn by the child one day per week for 12 weeks. The device measures the number of words the child hears, differentiating between adult and child voice and electronic sounds. The data is analysed by software developed by LENA and then

informs guided conversations about when and how to improve communications. Home-Start home visiting volunteers have been specially trained to work with the pedometer and resulting graphs, sharing them with the family in a positive and accessible way each week. The training guidelines may be found in Appendix D.

Home Start UK is currently piloting LENA HOME™ using trained volunteers in four sites in the UK namely Oldham, Southwark, Leeds and West Dorset, as located in the map below; it is the data from these four sites which is used in this report.



LENA Home UK pilot sites

What is the LENA HOME-START project?

The project involves using feedback from a number of 'wearable tech' vests developed by LENA, which measure the number of spoken words a child hears in a day with the DLP. The DLP and vest does not analyse the content of conversations but acts as a 'word counter' and can differentiate between adult and child speech, as well as between human and digital conversation. Exploring the data collected when a child has been wearing the vest for 24hrs, trained staff and volunteers can help parents to improve talk and conversation with their children.

Theory of change

The LENA project has an impact on parents, children and volunteer coaches. The foundational belief of the LENA project is that all parents have the ability to improve their children's futures with the right tools and support, and that by supporting parents in improving children's language environment during their early brain development, there is long term impact on children's readiness for nursery and primary school. Volunteers' knowledge and skills increases through their training in the project.

The theory of change was developed with the pilot local Home-Starts as part of the induction to the project.

> ***The proposition is that parents who are supported by trained volunteers to become more aware of the influences of a positive home learning environment and the importance of their vocal interactions with their children will change their beliefs, family routines and behaviours. These changes in parental beliefs, behaviours and family routines lead to improved outcomes for children specifically around early language acquisition. Therefore, interventions should not be exclusively focused on the child but should aim to improve interactions between children, parents, peers and the child's wider adult network.***

Families participating in the LENA project are paired with a trained volunteer coach who meets them each week to explain how they can increase the number of words the child hears, tailoring their support to the family's individual needs and using the word count and conversation data as both a springboard and a measure of progress. Coaches modelled positive engagement. A structured age appropriate curriculum provides sample activities to inform the quantity and quality of parent-child interactions –see appendix D.

Each pilot site had to meet a minimum IT specification to support the correct management of the data recorders and production of the resulting graphs. This IT specification has been developed and amended as part of the pilot process.

METHOD

Participants

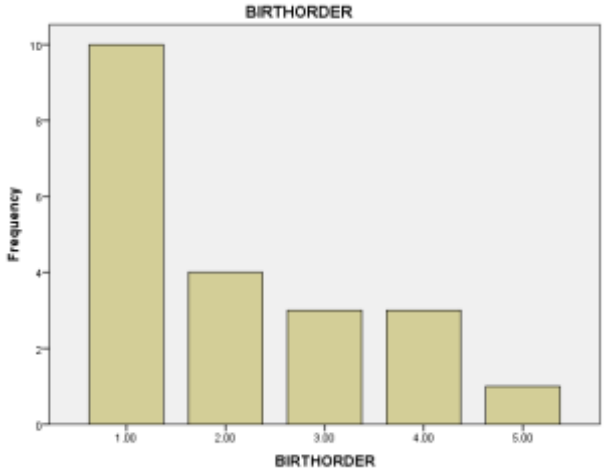
All the LENA families were taken from Home-Start families with a complex range of needs. This means they had already been identified as requiring support, possibly across a number of developmental areas. The families supported were reported by the local Home-Starts as being complex. Families engaged on the LENA project are varied in their backgrounds, socio-economic group, ethnicity, and make-up. However, as 3 of the four pilot sites are inner city schemes with existing case-loads of families in need, poorer families have made up the bulk of the numbers. Asylum seeking families are common in the numbers from Home-Start Oldham, Stockport and Tameside (HOST), and all families from Leeds are on the child protection or child in need register.

Thirty-five UK based families registered for the LENA programme. Of these there were 21 families with full data sets, i.e. with 12 recordings, pre and post assessment, snapshots and MESH scores. In addition, there were approximately 13 with pre-assessment and partial records on the other time points (those who dropped out part way through). Of these 21 families, the mean age of the children was 19.05 months with a range of 4 to 34 months. The modal gender was male but there were comparable numbers of boys and girls (11 boys/10 girls) and the modal birth order was first born although the range was 1-5th position and the median position was second. Home-Start family records show 9 of the families are lone parents, 6 have experienced domestic abuse and 17 report mental health issues. Four of the families are recorded as having adult learning needs. No further analysis was carried out of the children with incomplete datasets although clearly there is a question as to whether they exhibited distinctive characteristics which would enhance interpretation of the results.

In the control group, we also have 21 children as taken from the LENA normative database in the US. These children were individually matched on age, gender, birth order, and mother's education to the UK intervention group.

Figure 1 below shows the distribution of birth order in the UK Home-Start sample.

Figure 1: Distribution of birth order in the UK Home-Start sample



The range of maternal education was from no GCSEs through to Master's degree with the modal maternal education being no GSCEs but with the median being Diploma level education. One of the key descriptors was maternal education and this was coded as follows. The question was "What is the highest educational level obtained by the child's mother?" with the equivalent scoring for the UK

and we then compared the UK and the US scoring for maternal education so that the two were compatible. Figure 2 below shows the distribution of maternal education in the sample. We agreed the three-way classification shown in Table 1 with colleagues at the LENA Foundation.

Figure 2: Distribution of maternal education in the UK Home-Start sample.

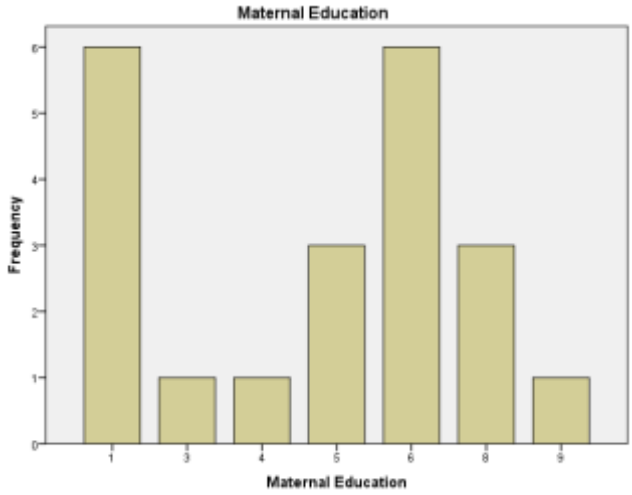


Table 1 Comparison between maternal education levels in UK and US samples.

The majority of the children (15/21) were from monolingual backgrounds

In the following pages we look at the results on the measures identified for the 21 children who went through the LENA Home-Start intervention. In each case we follow the same pattern looking at the HOME inventory the Preschool Language Scales and then the LENA measures. This pattern is then repeated with the matched US children who we have used as our control group. We report the results using the appropriate statistics and providing data tables throughout to give the reader the opportunity to look at the data.

There were no statistical differences in the language scores or responses for gender, birth order, maternal education or bilingual background, only significant differences related to Maternal education and the HOME (see Table 2below) where two of the subscales (Responsivity and Learning materials) and the Total HOME score showed significant differences and Bonferroni Corrections showed that each of the three maternal education groups (indicated in Table 2 above) were different from the other, with the highest scores in the households where the mothers had the highest education.

Table 2: HOME Inventory scores related to three levels of maternal education

		Sum of Squares	df	Mean Square	F	Sig.
Responsivity	Between Groups	56.921	2	28.461	7.416	**.004**
	Within Groups	69.079	18	3.838		
	Total	126.000	20			
Acceptance	Between Groups	6.674	2	3.337	1.656	.219
	Within Groups	36.279	18	2.015		
	Total	42.952	20			
Organisation	Between Groups	3.131	2	1.565	2.327	.126
	Within Groups	12.107	18	.673		
	Total	15.238	20			

Learning Materials	Between Groups	32.945	2	16.473	8.043	**.003**
	Within Groups	36.864	18	2.048		
	Total	69.810	20			
Involvement	Between Groups	15.124	2	7.562	2.426	.117
	Within Groups	56.114	18	3.117		
	Total	71.238	20			
Variety	Between Groups	3.588	2	1.794	1.180	.330
	Within Groups	27.364	18	1.520		
	Total	30.952	20			
Total	Between Groups	518.793	2	259.396	9.456	**.002**
	Within Groups	493.779	18	27.432		
	Total	1012.571	20			

Volunteers

Volunteers were recruited and trained locally in cohorts. The HSUK led 'train the trainer' system which was run according to local demand. The backgrounds and experience of volunteers to date has been varied, including early years, teaching, children's centre work, social work, experience of adoption and fostering, community work, nursing, teaching assistants, health visiting, speech and language therapist, midwife, nanny, doula, young mothers and grandparents.

Sustainability

Home-Start UK have developed the manualisation of the project, developed a business model to underpin its sustainability and provided opportunities for developing a funding proposition around LENA. In addition, the IT specification and requirements for the LENA project and been developed to ensure full accountability and compliance with GDPR.

The proposal was originally for 15 children in each of 4 geographical sites (i.e. 60 in total). The training of the volunteer coaches and the "local dimension" is not part of the evaluation.

The evaluation was three phased as follows:

Phase 1 SETUP

The collection of basic demographic data, the HOME inventory, and the Preschool Language Scale (Zimmerman et al, 2011). The children were then seen on a weekly basis in their homes for the duration of the curriculum. In practice children entered the programme over a period of time. The first children were recruited in the summer of 2017 and the last assessments were complete in early 2018

Phase 2 FOLLOW UP LENA MEETING

At the completion of the curriculum all families were visited again and Home and PLS repeated. LENA data was downloaded for the time period and again compared with the LENA database to give some sense of a control group. These measures will be carried out by someone who was not involved in the intervention to reduce potential bias.

Phase 3 DATA ANALYSIS

As a comparison, data from the US-based LENA data normative corpus with a sample of children, matched at the individual level for age of child, maternal education, gender and parity. This normative database provided by the LENA Research Foundation is compiled of more than 38,000 hours of spontaneous speech data that took place between January 2006 and December 2008. The speech recording data is stored in the LENA Natural Language Corpus, from which LENA derived normative information for the Adult Word Count estimates (AWC; adult words spoken per day), Conversational Turns estimates (CT; adult-child alternations per day), and Child Vocalization frequency estimates (CV; words, babbles, and pre-speech communicative sounds) and AVA that are reported in the LENA System, (Gilkerson and Richards 2008).

The data from the LENA Home-Start project children and their families together with LENA data from matched children (control data) were anonymised, by Home Start, and transferred to Newcastle University by June 2018. The use of the LENA was piloted up to March 2018.

The measures

The evaluation included the following data collection:

- Children's tested language performance using standardised measures,
- Collection of LENA data
- Comparison of LENA data with US-based LENA database which acted as 'control' data – i.e. matched children not receiving any intervention but followed up over time

The Preschool Language Scale PLS-UK

Preschool Language Scale-5 - UK (Zimmerman et al.2014) is a comprehensive developmental language assessment, with items that range from pre-verbal, interaction-based skills to emerging language and early literacy. The PLS-5 is designed for use with children aged birth through 7;11 to assess language development and identify children who have a language delay or disorder. The test aims to identify receptive and expressive language skills in the areas of attention, gesture, play, vocal development, social communication, vocabulary, concepts, language structure, integrative language, and emergent literacy (Examiner's Manual, pg. 3). The PLS-5 consists of two standardized scales: Auditory Comprehension (AC), to "evaluate the scope of a child's comprehension of language," and Expressive Communication (EC), to "determine how well a child communicates with others"(Examiner's Manual, pg. 4). Administration time varies based on the child's age and can range between 25-35 minutes for children aged birth through 11 months to 45-60 minutes for children over one year. Specific AC tasks assessed include comprehension of basic vocabulary, concepts, morphology, syntax, comparisons and inferences, and emergent literacy. Specific EC skills include naming, describing, expressing quantity, using specific prepositions, grammatical markers, sentence structures, and emergent literacy skills. Three optional supplemental measures are also included (Language Sample Checklist, Articulation Screener, and Home Communication Questionnaire). Scores are provided at three-month intervals from birth through 11 months, and at 6 months intervals

from 1 year through 7;11. The PLS-5 yields norm-referenced scores including standard scores, percentile ranks and age equivalents for the AC and EC scales as well as for Total Language (TL).

The HOME

The child's environment was assessed using the HOME Inventory (Caldwell and Bradley, 2003). This is designed to measure the quality and quantity of stimulation and support available to a child in the home environment. HOME is an acronym for Home Observation for Measurement of the Environment. Different versions are available for use depending on the age of the child; the Infant/Toddler version was appropriate throughout the LENA project. The HOME Inventory consists of 45 observable items divided into six areas; Responsivity, Acceptance, Organization, Learning Materials, Involvement and Variety. The observer, in this case the Home-Start UK Early Years Case Manager, is required to mark each point as observed or not. Examples include "Parent tells child name of object or provides toys that challenge child to develop new skills" and "Child has 3 or more books of his/her own".

LENA Assessment data

The Developmental Snapshot is a statistically validated, 52-item, parent completed evaluation of language skills for infants and toddlers focusing on well-established milestones associated with expressive and receptive language skills. The Snapshot is an assessment tool designed to gauge the child's language developmental age and language skills. The Snapshot consists of a series of questions that the parent answers about their child's expressive and receptive language skills. It is completed as frequently as once per month and provides developmental age and percentile ranking information compared to age-matched peers

The LENA technology, as described, analyzes automatically identified adult male and adult female segments and key child voice segments to generate estimates of the three environmental behavioral measures used in the current study: (a) Adult Word Count (**AWC**),) Conversational Turn-taking Counts (**CTC**) (Gilkerson et al. 2017) AWC is an estimate of the number of adult words spoken loudly enough to register clearly in the LENA recorder, but it does not differentiate child-directed speech from overheard speech. In practice, we estimate these words typically occur within a 10-ft radius of the child wearing the recorder. The child vocalization count (**CVC**) reflects the number of speech-related vocalizations produced by the child as identified by the automated procedure. For example, because child vocalizations are separated by 300 ms of silence and may be of varying length, the vowel "a" spoken in isolation or the babble "babababa," with no pauses or breaks, would both be assigned a count of one vocalization. Likewise, the string "mommy I want a cookie" would also count as one vocalization, provided that there was not more than 300 ms of silence between words or syllables. CT counts are the number of alternations within a conversation between clear, speech-related adult and key child vocalizations, as labeled by the automated procedure. A conversation was defined as a sequence of vocalizations bounded by at least 5 s of nonvocal material, based, in part, on rules suggested by Hart and Risley (1995). In this formulation, either child or adult may initiate a turn, and responses may not serve as the initiation of a subsequent turn. Thus, both of the sequences child–adult and child–adult–child, are counted as one and only one turn. If the parent or child interrupts the initiator, as is often the case in spontaneous speech interactions, the system will identify that section as overlapping speech, but the vocalization segment immediately following the overlap segment will be coded as a turn (given that the overlap section is not more than 5 s.

ETHICAL CONSIDERATIONS

Ethical approval for the project was organised by Home-Start at site level.

ANALYTIC PLAN

We then carried out a series of analyses. For clarity they are separated out into within group changes and between group changes. 'Within group' means how much the children in our intervention group changed on the Language, HOME and LENA measures over the period that they were involved in the study. For the language and HOME measures this was a six-month period. For the LENA data this was over a three-month (twelve week) period. The data were tested for normality and where appropriate parametric and non-parametric tests were used. The key questions here are do the specific aspects tested change statistically between Time 1 (the start) and Time 2 (the end) of the intervention period? So, does our intervention group change on the measures concerned.

The 'between group' analysis involves comparing our intervention data with that from the LENA Foundation matched control group. The way that the LENA group data were presented was 4 recordings approximately 4-weeks apart. Therefore, to match this data to our own for comparison, we took from our data the recordings from week 1, week 5 and week 9 (all 4 weeks apart) and compared it to their first 3 time points of data. To make things clearer in the Table 3 below we indicate which analyses were within and which were between groups. So, this approach answers the question "do the intervention group on average have a higher or a lower score that matched controls would have had over the same time period. To make it easier to read we have then numbered the analyses so that these can be tracked below (see Table 3). The specific statistics used are provided in the text below, as are the descriptive data for the specific variables. For example, there are time plots for the LENA variables and these are provided in the text but are not separate analyses.

As part of the analysis we have also looked at the associations between different variables to see whether they are closely associated and ask questions such as "if parents speak more to their children are there more conversational turns?" or are they just speaking more? We have also included regression models to explore the predictive relationship between our key demographic data to which reference was made above and the outcome.

Table 3: Variables by types of analyses

Analyses and Variables	**Within group analyses**	**Between group analyses**	**Analysis #**
HOME			
Home Data; Pre- and post-intervention associations	X		1
Home Subscale Change:			
Responsivity	X		2
Acceptance	X		3
Organisation	X		4
Learning Materials	X		5
Involvement	X		6
Variety	X		7
Total	X		8
HOME Regression	X		9
LANGUAGE			
Language Test Data; Pre- and post-intervention scores	X		10
Language Test Data; Difference between time points:			
Comprehension raw score	X		11
Comprehension standard score	X		12
Comprehension percentile score	X		13
Expressive raw score	X		14
Expressive standard score	X		15
Expressive percentile score	X		16

Language Total raw score	X		17
Language Total standard score	X		18
Language Total percentile score	X		19
Language Test Regression	X		20
LENA			
Developmental Snapshot (Age)	X		21
Boxplots	X		22
Descriptive Statistics	X		23
Change Over Time; Adult Word Count (AWC)	X	X	24
Change Over Time; Conversational Turns (CT)	X	X	25
Change Over Time; Child Vocalisation Count (CV)	X	X	26
Change Over Time; Automatic Vocalisation Assessment (AVA)	X		27
Change Over Time; TV (Secs)	X		28
LENA Case Studies	X		29
The Comparison Group			
Adult Word Count (AWC)		X	30
Child Vocalisation Count (CV)		X	31
Conversational Turns (CT)		X	32
Correlations Between Variables	X	X	33
LENA Regression	X		34

As can be seen from the descriptive portrayals of the LENA data there was considerable variation in many of the analyses from week to week. Accordingly, for the LENA data only we sought to ask a specific question about whether on average children's score changed in an upwards direction from week 1 to an average of week 2-12. We denoted this as Average Change.

FINDINGS

ANALYSIS 1: HOME Data; Pre- and Post-intervention associations

Initially we look at the relationship between the HOME score subscales and Total score at the two time points, immediately before and immediately after intervention. As we can see that there is a good deal of association across time as might be expected. Some of these associations are statistically significant but there is also a lot of difference on these subscales over a relatively short time period. Where the correlations are lower there is clearly more change happening between the two time points. It is important to acknowledge that there are no examples of negative correlations with HOME scores going down over time.

Table 4: Correlations between Pre- and Post- intervention Home scores

		Responsivity	Acceptance	Organisation	Learning Materials	Involvement	Variety	Total
Spearman's rho	Responsivity	.657	.472	.317	.227	.230	.486	.606
		.001	**.031**	.161	.322	.316	**.025**	**.004**
		21	21	21	21	21	21	21
	Acceptance	.501	.772	.128	.290	.533	.140	.711
		.021	**.000**	.580	.203	**.013**	.545	**.000**
		21	21	21	21	21	21	21
	Organisation	.700	.296	.613	.307	.162	.183	.631
		.000	.193	**.003**	.176	.484	.426	**.002**
		21	21	21	21	21	21	21
	Learning Materials	.483	.578	-.009	.388	.280	.501	.549
		.027	**.006**	.971	.082	.219	**.021**	**.010**
		21	21	21	21	21	21	21
	Involvement	.166	.282	-.112	.024	.301	.442	.210
		.473	.215	.630	.919	.185	**.045**	.361
		21	21	21	21	21	21	21
	Variety	-.009	.213	.186	.252	-.013	.686	.262
		.968	.355	.420	.271	.956	**.001**	.251
		21	21	21	21	21	21	21
	Total	.588	.666	.191	.375	.364	.588	.711
		.005	**.001**	.406	.094	.104	**.005**	**.000**
		21	21	21	21	21	21	21

We then turn to the distributions of each of the HOME subscales for each time points.

Table 5: Time 1 distribution of HOME scores

	Responsivity	Acceptance	Organisation	Learning Materials	Involvement	Variety	Total
Median	10.0000	6.0000	4.0000	7.0000	3.0000	3.0000	33.0000
Range	8.00	6.00	3.00	7.00	6.00	4.00	31.00
Minimum	3.00	2.00	3.00	1.00	.00	1.00	12.00
Maximum	11.00	8.00	6.00	8.00	6.00	5.00	43.00

Table 6: Time 2 distribution of HOME scores

	Responsivity	Acceptance	Organisation	Learning Materials	Involvement	Variety	Total
Median	11.0000	6.0000	5.0000	8.0000	4.0000	4.0000	37.0000
Range	6.00	6.00	2.00	9.00	5.00	3.00	24.00
Minimum	5.00	2.00	4.00	.00	1.00	2.00	19.00
Maximum	11.00	8.00	6.00	9.00	6.00	5.00	43.00

In each case we see that, in all but one scale, the median has increased across the two times points. If we then test this we see that in a number of instances these changes are statistically significant.

ANALYSES 2-8: HOME Subscale Change

Table 7: Significance of change in subscale scores from pre- to post-intervention

HOME subscale and Total scores compared between Time 1 and Time 2[a]

	Responsivity	Acceptance	Organisation	Learning Materials	Involvement	Variety	Total
Z	-1.987[b]	-1.807[b]	-2.138[b]	-2.830[b]	-.991[b]	-3.082[b]	-3.427[b]
Asymp. Sig. (2-tailed)	**.047**	**.071**	**.033**	**.005**	.322	**.002**	**.001**

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

These data would suggest that during the course of the intervention 5 of the six subscales plus the total score changed significantly. The suggestion here would be that the type of intervention had a direct bearing on the home learning environment in a number of interesting ways. Of course, care has to be taken not to over interpret these findings because without experimental control we cannot claim causation here. But by the same token there is no reason to assume that these scores would change without other factors in the environment having an effect. It is also important that these HOME measures were not carried out blind to the status of the children or families, but they were carried out by the external researcher to the project not by the person delivering the intervention.

ANALYSIS 9: HOME Regression

Regression analysis allows us to explore which background characteristics of families and children are the strongest predictors of our outcome data.

Predictor variables included in the regression analysis for the HOME outcomes were:

- Child Age (months)

- Gender (Male/Female)
- Birth Order (1,2,3,4,5 etc)
- Maternal Education
- Bilingual child (Yes/No)

It would, of course, have been possible to make these models more complex including for example structural measures of social disadvantage such as income or housing and variables such as maternal mental health which have been shown in other studies to predictive value. However, it was agreed, at the outset, that such questions should be as un-invasive as possible to facilitate family engagement in the intervention.

Responsivity

Correlation analysis revealed significant positive correlations between the Responsivity scale on the HOME and maternal education ($r=.630$, $p<0.01$), and whether the child was bilingual ($r=.387$, $p<0.05$).

Predictor variables altogether accounted for 44% of the total variance in Responsivity scores. The regression model was not significant ($p>0.05$) however in-line with the correlation analysis the strongest predictor of Responsivity was maternal education ($t(15)=2.726$, $p<0.05$).

Acceptance

Correlation analysis indicated significant negative correlation between the Acceptance scale on the HOME and child age ($r=-.378$, $p<0.05$) (Acceptance score decreased as age increased), and significant positive correlation between Acceptance and maternal education ($r=.382$, $p<0.05$), and whether the child is bilingual ($r=.537$, $p<0.01$).

Predictor variables altogether accounted for 47% of the total variance in Acceptance scores, but the regression model was non-significant ($p>0.05$).

Organisation

Correlation analysis indicated significant positive correlation between Organisation scores on the HOME and maternal education ($r=.439$, $p<0.05$).

Predictor variables altogether accounted for only 28% of the total variance on Organisation score and the regression model was non-significant ($p<0.05$). Maternal education was the strongest predictor of Organisation score ($t(15)=2.360$, $p<0.05$).

Learning Materials

Correlation analysis revealed significant positive correlations between the Learning Materials scale on the HOME and maternal education ($r=.673$, $p<0.01$), and whether the child was bilingual ($r=.496$, $p<0.05$).

Predictor variables altogether accounted for 56% of the total variance in Learning Materials scores. The regression model was significant ($F(5,15)=3.825$, $p<0.05$). The strongest predictor of Learning Materials was maternal education ($t(15)=2.788$, $p<0.05$).

Involvement

Correlation analysis revealed significant positive correlations between the Involvement scale on the HOME and maternal education (r=.458, p<0.05),

> *"My son seems a lot calmer and I think that's because I am taking more time to play with him"*
>
> *Mum 2073*

Predictor variables altogether accounted for 25% of the total variance in Involvement scores. The regression model was non- significant (p>0.05).

Variety

Correlation analysis indicated a significant positive correlation between Variety score on the HOME and birth order (r=405, p<0.05).

Predictor variables altogether accounted for 28% of the total variance of Variety score, and the regression model was non-significant (p>0.05).

Total HOME Score

Correlation analysis revealed significant positive correlations between the Total HOME score and maternal education (r=.710, p<0.01), and whether the child was bilingual (r=.499, p<0.05).

Predictor variables altogether accounted for 58% of the total variance in Total HOME scores. The regression model was significant (F(5,15)=4.274, p<0.05). The strongest predictor of Total HOME scores was maternal education (t(15)=3.423, p<0.01).

ANALYSIS 10: Language Test Data; Pre- and Post-intervention scores

If we start by looking at the association between the specific subscale scores of the preschool Language Scale at the start and the end of the intervention.

Table 8: Correlations between pre- and post-intervention Language test scores

	PLS Comprehension Raw Score	PLS Comprehension Standard Score	PLS Comprehension Percentile Rank	PLS Expression Raw Score	PLS Expression Standard Score	PLS Expression Percentile Rank	PLS Total Raw Score	PLS Total Standard Score	PLS Total Percentile Rank
PLS Comprehension Raw Score	.901 **.000** 21	-.209 .363 21	-.325 .151 21	.909 **.000** 21	.024 .919 21	.031 .894 21	-.101 .662 21	-.121 .601 21	-.155 .501 21
PLS Comprehension Standard Score	-.331 .142 21	.693 **.001** 21	.648 **.001** 21	-.322 .155 21	.667 **.001** 21	.624 **.003** 21	.713 **.000** 21	.718 **.000** 21	.682 **.001** 21
PLS Comprehension Percentile Rank	-.382 .088 21	.579 **.006** 21	.649 **.001** 21	-.335 .136 21	.604 **.004** 21	.606 **.004** 21	.624 **.002** 21	.631 **.002** 21	.681 **.001** 21
PLS Expression Raw Score	.728 **.000** 21	-.364 .104 21	-.411 .064 21	.796 **.000** 21	-.081 .727 21	-.062 .788 21	-.238 .298 21	-.257 .261 21	-.255 .265 21
PLS Expression Standard Score	-.093 .688 21	.205 .373 21	.242 .290 21	-.037 .876 21	.344 .127 21	.329 .146 21	.287 .208 21	.294 .212 21	.298 .189 21
PLS Expression Percentile Rank	-.212 .356 21	.210 .361 21	.255 .264 21	-.085 .715 21	.394 .077 21	.385 .085 21	.318 .163 21	.317 .162 21	.340 .132 21
PLS Total Raw Score	-.235 .306 21	.492 **.023** 21	.496 **.022** 21	-.177 .444 21	.589 **.007** 21	.536 **.012** 21	.559 **.008** 21	.559 **.008** 21	.549 **.010** 21
PLS Total Standard Score	-.240 .294 21	.506 **.019** 21	.511 **.018** 21	-.193 .426 21	.579 **.006** 21	.548 **.010** 21	.571 **.007** 21	.571 **.007** 21	.564 **.008** 21
PLS Total Percentile Rank	-.313 .167 21	.471 **.031** 21	.545 **.011** 21	-.213 .354 21	.597 **.004** 21	.602 **.004** 21	.562 **.008** 21	.567 **.007** 21	.621 **.003** 21

The first thing to notice is that unsurprisingly many of these variables are highly associated across time, especially with regard to the raw score. The higher you start the higher you end up. The unusual exceptions are the standard score and percentile ranks for the expressive scale which suggest that this is rather less stable than the other scales (and thus potentially more amenable to change).

Table: 9 Time 1 distribution of PLS Scores

Time 1 PLS Scores

	PLS Comp Raw Score	PLS Comp Standard Score	PLS Comp Percent. Rank	PLS Exp Raw Score	PLS Exp Standard Score	PLS Exp Percent. Rank	PLS Total Raw Score	PLS Total Standard Score	PLS Total Percent. Rank
Mean	18.19	85.86	21.38	18.9524	85.09	22.38	170.95	84.57	20.04
Std. Deviation	4.308	10.98	20.29	5.73128	12.63	21.11	20.51	10.77	18.88
Minimum	9.00	63.00	1.00	7.00	63.00	1.00	140.00	68.00	2.00
Maximum	26.00	116.00	86.00	28.00	109.00	73.00	215.00	108.00	70.00

We then turn to the distribution of the tests scores at each time point, starting with Time 1.

Table 9: Time 1 distribution of PLS scores

Time 2 PLS Scores

	PLS Comp Raw Score	PLS Comp Standard Score	PLS Comp Percent. Rank	PLS Exp Raw Score	PLS Exp Standard Score	PLS Exp Percent. Rank	PLS Total Raw Score	PLS Total Standard Score	PLS Total Percent. Rank
Mean	22.19	85.57	21.48	24.48	90.67	29.76	176.24	87.33	24.14
Std. Deviation	5.33	10.62	16.48	5.58	10.04	20.60	19.57	10.37	18.41
Minimum	12.00	57.00	2.00	10.00	71.00	3.00	128.00	62.00	1.00
Maximum	33.00	102.00	55.00	35.00	113.00	81.00	213.00	107.00	68.00

And then look at the time 2 test scores

ANALYSES 11-19: Language Test Data; Difference between time points

And then we test the difference between the two using ANOVA.

Table 10: Pre- and post-intervention differences in PLS scores

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PLS Comprehension Raw Score - PLS Comprehension Raw Score	-4.00000	2.36643	.51640	-5.07719	-2.92281	-7.746	20	**.000**
Pair 2	PLS Comprehension Standard Score - PLS Comprehension Standard Score	.28571	8.60316	1.87736	-3.63040	4.20182	.152	20	.881
Pair 3	PLS Comprehension Percentile Rank - PLS Comprehension Percentile Rank	-.09524	15.80160	3.44819	-7.28804	7.09756	-.028	20	.978
Pair 4	PLS Expression Raw Score - PLS Expression Raw Score	-5.52381	3.61413	.78867	-7.16894	-3.87868	-7.004	20	**.000**
Pair 5	PLS Expression Standard Score - PLS Expression Standard Score	-5.57143	13.16272	2.87234	-11.56303	.42017	-1.940	20	.067
Pair 6	PLS Expression Percentile Rank - PLS Expression Percentile Rank	-7.38095	23.12677	5.04668	-17.90813	3.14623	-1.463	20	.159
Pair 7	PLS Total Raw Score - PLS Total Raw Score	-5.28571	18.84978	4.11336	-13.86603	3.29460	-1.285	20	.213
Pair 8	PLS Total Standard Score - PLS Total Standard Score	-2.76190	9.79747	2.13798	-7.22166	1.69785	-1.292	20	.211
Pair 9	PLS Total Percentile Rank - PLS Total Percentile Rank	-4.09524	16.24163	3.54421	-11.48834	3.29786	-1.155	20	.262

From this we conclude that the expressive and the receptive raw scores have increased significantly whereas the raw score for the Total PLS score, rather surprisingly, has not. By contrast the standard scores have not increased significantly suggesting that the changes are not enough to change the order of the child relative to the standardisation sample of the test. Their scores have not gone up more than those of other children.

ANALYSIS 20: Language Test Regression

As we have seen regression analysis allows us to explore which background characteristics of families and children are the strongest predictors of our outcome data. Predictor variables included in this regression analysis for the Language Test outcomes were:

- Child Age (months)
- Gender (Male/Female)
- Birth Order (1,2,3,4,5 etc)
- Maternal Education
- Bilingual child (Yes/No)

PLS Comprehension Raw Score

Correlation analysis revealed significant positive correlations between PLS Comprehension raw score and child age ($r=.773$, $p<0.01$), and birth order ($r=.379$, $p<0.05$). Predictor variables altogether accounted for 73% of the variance in PLS Comprehension score ($r2=.737$). The regression model was

significant (F(5, 15)=8.420, p<0.01). As expected, child age was the strongest predictor of PLS Comprehension Score (t(15)=5.212, p<0.01).

PLS Expression Raw Score

Correlation analysis indicated significant positive correlations between PLS Expression raw score and child age (r=.801, p<0.01). Predictor variables altogether accounted for 67% of the variance in PLS Expression score (r2=.671). The regression model was significant (F(5,15)=6.130, P<0.01). As expected, child age was the strongest predictor of PLS Expression Score (t(15)=4.958, p<0.01).

ANALYSIS 21 LENA data; Developmental Snapshot

We now turn to the LENA data and start by looking at the LENA Developmental Snapshot data comparing performance between Time 1 and Time 3. All 21 children had data from these two points. Five children had this for four time points.

Unsurprisingly the two time points 1 and 3 were all correlated with one another.

Table 11: LENA Developmental Snapshot data paired comparison between T1 and T3.

		N	Correlation	Sig.
Pair 1	LENADSn1Age & LENADSn3Age	20	.846	**.000**
Pair 2	LENADSn1Score & LENADSn3Score	20	.848	**.000**
Pair 3	LENADSn1Ss & LENADSn3Ss	20	.743	**.000**
Pair 4	LENADSn1Pctl & LENADSn3Pctl	20	.801	**.000**

Table 12: Statistical significance of change in Developmental Snapshot data between T1 and T3

Pair	Variable	Mean	SD	t	df	p
1	Developmental Snapshot Age T1-T3	5.52	4.53	-5.637	20	.000
2	Developmental Snapshot Score T1-T3	6.95	5.49	-5.793	20	.000
3	Developmental Snapshot Standard Score	7.03	14.31	-2.251	20	.036
4	Developmental Snapshot Percentile	9.42	20.95	-2.062	20	0.052

"It has been interesting to see how more responsive my daughter has become"

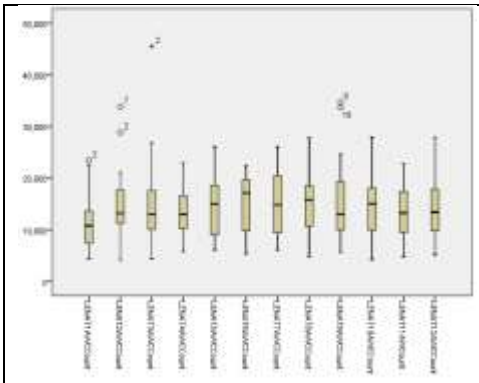
Mum 2198

If we compare the two time points statistically we see significant differences for the developmental snapshot age, score data and standard scores and marginal significant differences for percentile data. On all variables mean scores increased over time which indicates children's expressive and receptive language ability improved significantly over time.

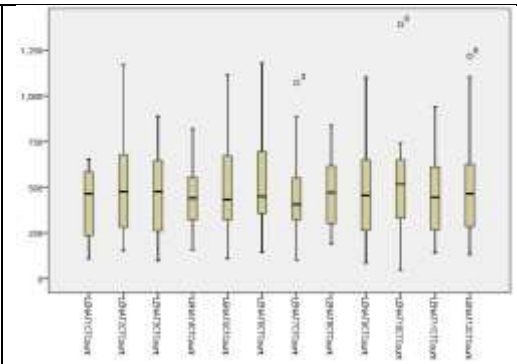
ANALYSES 22: Boxplots

We now turn to the data taken automatically from the body worn LENA recorders. Each recording is for one day per week. We have confined our analyses here to the 21 children identified above. 20 of the 21 children had all 12 recordings, and 1 had 11 recordings (missing the final 12th day). We start by looking at each of the four LENA outcomes: Adult Word Count, Conversational Turns, Automatic Vocalisation Assessment, and TV. In each case we report the count score but in the AVA we report a percentile score. The x axis reports the 12 recordings the y axis the LENA scores. Finally, for the TV sound which is an indication of how long the Television is on in the home during the day, the y axis is the number of seconds that the TV is on during the period concerned.

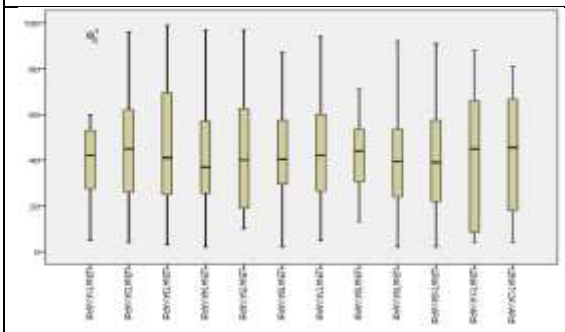
Figures 3-7: Boxplots of key LENA variables over 12 weeks



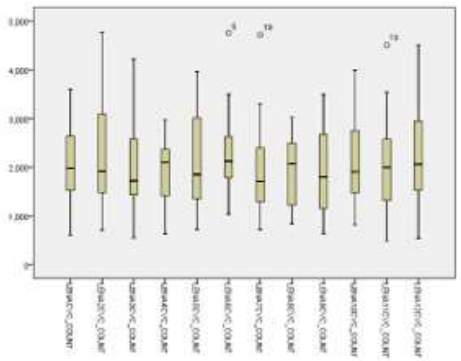
The Adult Word count



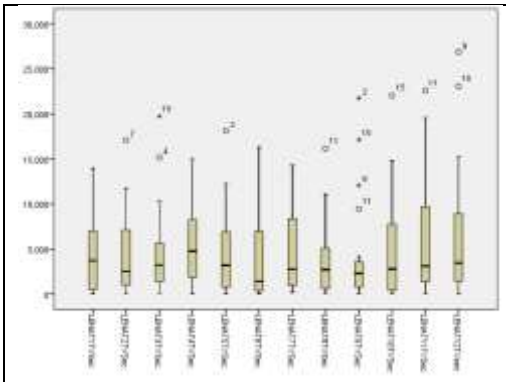
The Conversational Turns Count



Automatic Vocalisation Assessment Percentile



The Child Vocalisation



The TV Count

The box plots below show the variation in scores over the 12-week period. Numbered dots in each plot are outliers in the data. Each 'box' represents the middle 50% of scores. The plots show the amount of movement there is from week to week. Whereas one might expect either a gradual increase or a flat line if the noise in the distribution simply ended up with one child's profile cancelling out that of another child, in fact we have a wave like pattern with the group seeming to go up and then down. It is not possible to explain these changes from the data we have here, but one might suggest that this is associated with the volunteer input (perhaps in terms of the programme curriculum) or in terms of fluctuations in parent/child intervention fatigue. Individual participant graphs of AWC, CT and AVA data over the 12 recordings can be found in Appendix B and we can see that both adults and children start from very different positions. In some cases, we see a clear rise in AWC from T1 to T2 (OLD2216, WODF545) but, in most cases, the patterns are very variable.

ANALYSES 23: Descriptive Statistics

The tables below display the minimum and maximum scores, mean scores and standard deviations for each of the 4 LENA outcomes over the 12 weeks.

Table 13: Descriptive statistics for Adult Word Count over 12 weeks

	Minimum	Maximum	Mean	Std. Deviation
LENAT1AWCCount	4414	23404	11393.71	4558.909
LENAT2AWCCount	4182	33812	15000.00	7036.814
LENAT3AWCCount	4374	45506	14922.62	8736.095
LENAT4AWCCount	5802	22977	13575.33	4886.161
LENAT5AWCCount	6139	26065	14104.33	6086.434
LENAT6AWCCount	5382	22321	14805.43	5793.114
LENAT7AWCCount	6106	27526	15822.33	6459.548
LENAT8AWCCount	4735	27850	15370.43	5538.018
LENAT9AWCCount	5628	34695	16409.67	8902.708
LENAT10AWCCount	4251	27959	14381.05	5423.226
LENAT11AWCount	4828	22831	13439.38	5416.593
LENAT12AWCCount	5129	27856	14362.00	5918.450

Table 14: Descriptive statistics of Conversational Turn count scores over 12 weeks.

	Minimum	Maximum	Mean	Std. Deviation
LENAT1CTCount	107	655	410.86	150.545
LENAT2CTCount	153	1169	530.38	312.453
LENAT3CTCount	98	889	464.05	214.420
LENAT4CTCount	155	818	430.57	173.529
LENAT5CTCount	109	1114	486.62	275.361
LENAT6CTCount	144	1179	507.95	254.409
LENAT7CTCount	100	1071	466.14	244.514
LENAT8CTCount	191	841	474.52	175.217
LENAT9CTCount	85	1101	491.00	273.134
LENAT10CTCount	45	1390	510.86	276.768
LENAT11CTCount	141	942	454.76	222.786
LENAT12CTCount	130	1218	505.40	282.786

Table 15: Descriptive statistics of Child Vocalisation Count scores over 12 weeks.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
LENACVC_COUNT	20	609	3600	2034.15	739.155
LENA2CVC_COUNT	20	710	4768	2324.95	1174.138
LENA3CVC_COUNT	20	554	4218	2050.70	946.172
LENA4CVC_COUNT	20	632	2967	1949.45	646.104
LENA5CVC_COUNT	20	721	3962	2117.25	979.698
LENA6CVC_COUNT	20	1038	4759	2278.90	935.616
LENA7CVC_COUNT	20	723	4719	1938.75	937.067
LENA8CVC_COUNT	20	842	3024	1919.85	675.576
LENA9CVC_COUNT	20	636	3497	1926.50	931.461
LENA10CVC_COUNT	20	825	3992	2067.70	877.009
LENA11CVC_COUNT	20	492	4507	2064.30	1060.720
LENA12CVC_COUNT	20	544	4510	2234.50	1087.316
Valid N (listwise)	20				

Table 16: Descriptive statistics of Automatic Vocalisation Assessment percentile scores over 12 weeks.

	Minimum	Maximum	Mean	Std. Deviation
LENAT1AVAPctl	5	95	42.33	23.478
LENAT2AVAPctl	4	96	46.10	25.020
LENAT3AVAPctl	3	99	47.86	26.083
LENAT4AVAPctl	2	97	41.67	22.948
LENAT5AVAPctl	10	97	42.76	25.861
LENAT6AVAPctl	2	87	41.19	21.908
LENAT7AVAPctl	5	94	42.29	23.599
LENAT8AVAPctl	13	71	41.48	17.247
LENAT9AVAPctl	2	92	40.05	21.607
LENAT10AVAPctl	2	91	40.90	24.240
LENAT11AVAPctl	4	88	42.76	29.115
LENAT12AVAPctl	4	81	41.40	25.656

Table 17: Descriptive statistics of TV time in seconds over 12 weeks.

	Minimum	Maximum	Mean	Std. Deviation
LENAT1TVSec	53	13976	4084.14	4200.620
LENAT2TVSec	93	17072	4207.00	4704.796
LENAT3TVSec	85	19729	4367.90	5080.360
LENAT4TVSec	61	14975	4988.38	3796.884
LENAT5TVSec	57	18113	5015.14	4970.913
LENAT6TVSec	54	16272	4279.95	5239.732
LENAT7TVSec	175	14279	4609.67	4753.269
LENAT8TVSec	82	16159	3815.52	4243.332
LENAT9TVSec	103	21740	4262.71	5887.405
LENAT10TVSec	93	22031	4478.33	5605.956
LENAT11TVSec	86	22565	5705.10	6419.486
LENAT12TVsec	101	26884	6349.60	7574.952

We then tested the relationship between the scores week on week for each item using a repeated measure ANOVA. Perhaps unsurprisingly given the variable movement across weeks the repeated measures ANOVA did not show statistically significant differences, for all four variables $p>0.05$. Following this analysis, we adjusted our data to explore our key issue; whether children change from baseline assessment, therefore we then tested the differences between the Time 1 score and an aggregate of the Time2 – Time 12 scores, addressing the question do children's LENA scores "on average" rise from their starting point in the programme.

ANALYSES 24: Change Over Time: AWC

We can note the variability between different time points – i.e. there is not a clean upwards trajectory and we also note that there is huge variability at each time point. It is also noteworthy that the total TV time has also increased. We then tested the significance of change in scores over time. To do this we took the first week's scores (T1) and averaged the scores from week 2-12 (T2-T12) to provide a second score for comparison. By averaging weeks 2-12 we are ironing out the large variability and outliers in data across the weeks. We then tested the difference between the first recording and the averaged recording to see if change in count scores over time was significant. The table below shows the mean Adult Word Count (AWC) scores for the two time points plus their standard deviations.

Table 18: Mean Adult Word Count scores for baseline time point 1, and averaged time points 2-12.

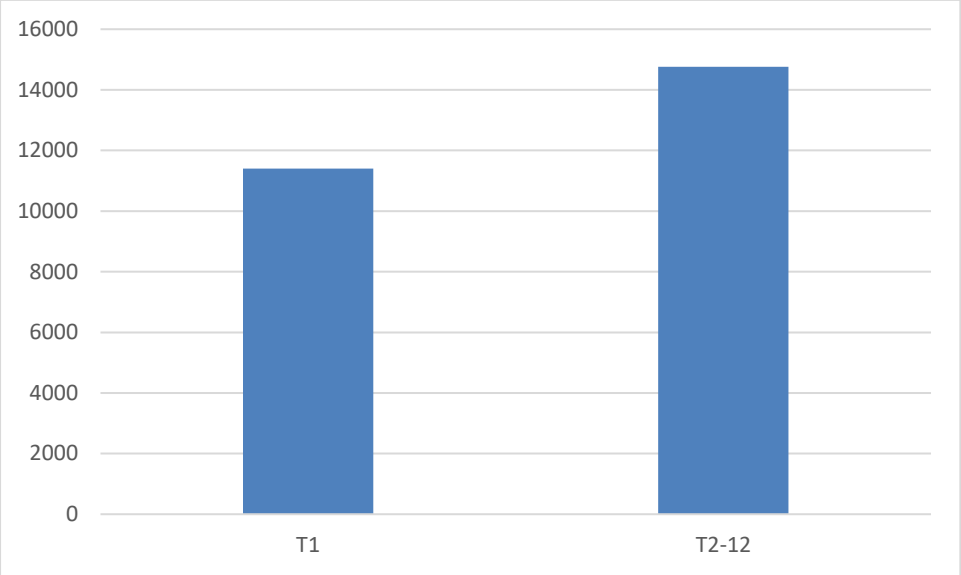
Variable	N	Mean	Standard deviation
AWC T1	21	11399.67	5174.128
AWC T2-12	21	14758.15	4404.669

Repeated measures analysis indicated a highly significant effect of time on AWC: $F(1,20) = 10.393$, $p=<0.01$. Therefore, there was a statistically significant increase in Adult Word Counts throughout the trial, this is also reflected in the figure below.

"I didn't realise how little I actually spoke to my child, this has really helped me understand how to support her development"

Mum 2455

Figure 8: Mean AWC score for T1 and T2-12.



ANALYSES 25: Change Over Time; CT

We then tested the difference in mean Conversational Turn scores over time. The table below shows the mean CT scores at baseline and over weeks 2-12.

Table 19: Mean Conversational Turn scores for baseline time point 1, and averaged time points 2-12.

Variable	N	Mean	Standard deviation
CT T1	21	439.33	175.00
CT T2-12	21	479.57	192.04

Repeated measures analysis indicated no significant effect of time on CT (p=>0.05). Therefore, the increase we see in the mean CT scores over time was not statistically significant.

ANALYSES 26: Change Over Time; CVC

"I can really see a difference with how much more my daughter talks compared to when my other children were this age"

Mum 1323

Table 20: Mean Child Vocalisation Count scores for baseline time point 1, and averaged time points 2-12.

Variable	N	Mean	Standard deviation
CVC T1	20	2034.15	739.15
CVC T2-12	20	2075.58	714.07

Repeated measures analysis indicated no significant effect of time on CVC (p=>0.05).

ANALYSES 27: Change Over Time; AVA

Automatic Vocalisation Assessment percentile scores were then tested for change over time. The table below displays the mean AVA percentile at baseline and over weeks 2-12.

Table 21: Mean Automatic Vocalisation Assessment percentile scores for baseline time point 1, and averaged time points 2-12.

Variable	N	Mean	Standard deviation
AVA T1	21	42.33	23.47
AVA T2-12	21	42.57	16.95

Repeated measures analysis indicated no significant effect of time on AVA (p=>0.05). As we see from mean scores at both time points, AVA remains in the 42nd percentile.

ANALYSES 28: Change Over Time; TV

Finally, we looked at the change in proportion of TV across time. The table below shows the mean TV time in seconds, at baseline, and over weeks 2-12.

Table 22: Mean TV time in seconds for baseline time point 1, and averaged time points 2-12.

Variable	N	Mean	Standard deviation
TV T1	21	4084.14	4200.62
TV T2-12	21	4716.76	3808.02

Repeated measures analysis indicated no significant effect of time on TV (p=>0.05) although it is important to note that the mean amount of TV time recorded had, like the other variables, risen.

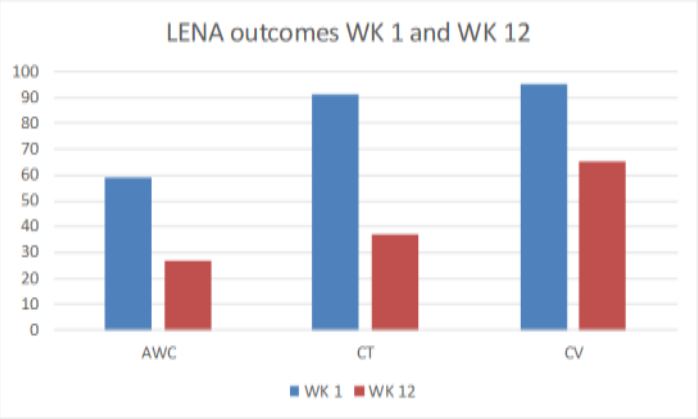
ANALYSIS 29: LENA Case Studies

Janey

Janey was born at 23 weeks,18 months ago. Her mother is single, this is her first child. She has no family locally and is an asylum seeker. English is not her first language, although she is quite confident in it. Janey's father has never seen his child. The first assessment booking was cancelled at short notice, as they were called to court regarding immigration.

Although the whole experience of having a premature baby has clearly had a huge impact on the mother, Janey's medical appointments show she doesn't seem to be suffering particularly as a result of her prematurity. The house is quite bare. *HSUK SLT*

Outcomes:

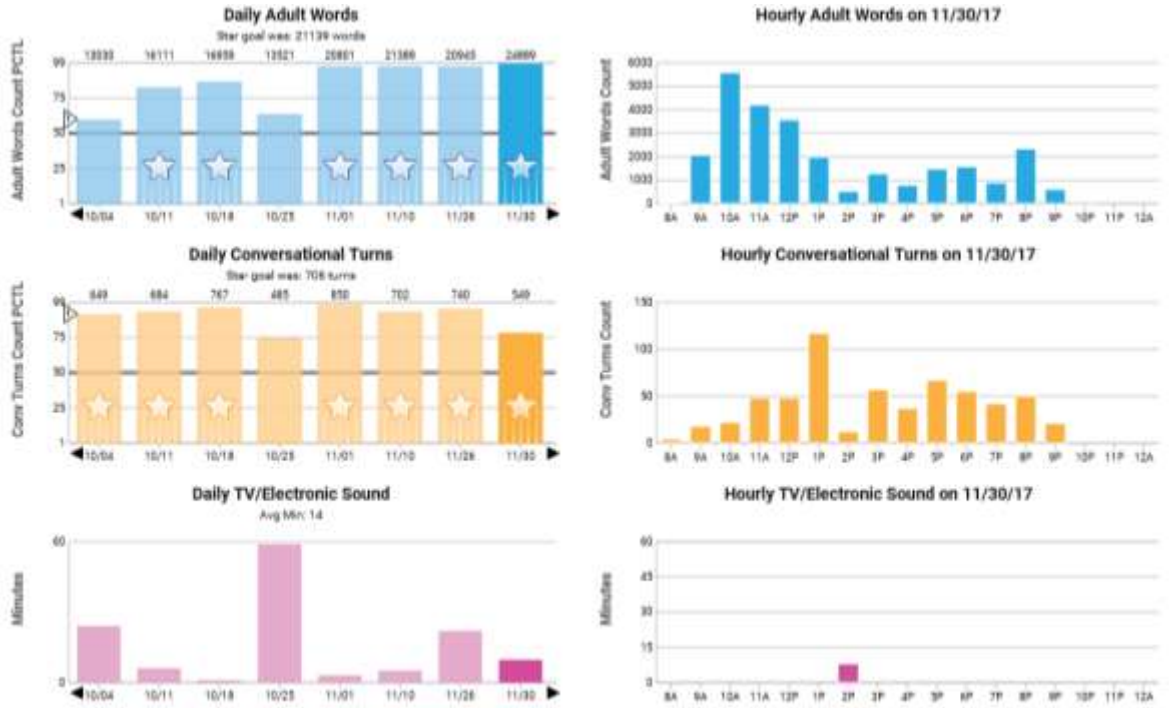


Data comparison between week 1 and week 12 above shows a decrease in all three areas of Adult Word Count, Conversational Turns and Child Vocalisations. However, looking at the information the family receive each week from their data through their volunteer shows a more detailed picture which built confidence and supported personal development across the weeks of the project.

The charts below are from the print out Janey and her mum received on week 8 of their 12 weeks. The printed charts which are taken to the family each week only have space for 8 weeks of data. These are Janey's first 8 weeks. On week 9 they received a print-out showing weeks 2-9.

On the left you can see three charts: daily adult words, daily conversational turns, and electronic sound in minutes. Each bar represents one day's recording data. The latest one is on the right of each chart. This means we can see all 8 weeks of progress here. The total number of words or turns is at the top of each bar. The date (in the American format) is at the bottom.

On the right is a breakdown of the data from the latest data of recording by hour. This facilitates personalised feedback from the volunteer to the family on specific activities.



Despite Janey's circumstances, her number of words heard is in the high average or high band from week one. Her conversational turns are almost always in the high band.

The LENA device does not record actual words, so the number of conversational turns are an important marker of actual interaction beyond simply hearing words. We can see that Janey is often being spoken to and responding – a conversational turn. Being able to recognise your own place in a conversation is also a key developmental skill. As the device records baby vocalisations too, this count is valid with all the children in the target group right from birth. A skilled volunteer will combine this with observation from their own interactions with/of mum and child to establish a complete picture.

An interesting note on these graphs is possible correlation in week 4 between lower than usual word count and lower than usual conversational turns with a higher than usual amount of TV/Electronic sound. In week 4 Janey still heard less than 60 minutes of TV/Electronic sound, but her other word counts were down. Again, the Home-Start LENA volunteer would explore explanations of this with the family to help them understand their interactions and impacts of activities.

The volunteer is supported in exploring what's happening by the hourly breakdowns on the right. The family are encouraged to keep a diary or informal note of what they did on the day the word-pedometer was worn, so that they can talk about it with their volunteer at the next visit. The visit where these graphs are discussed may include things like – 'Janey heard lots of words on this morning, that was great! Although she heard lots of words around 10am, we can see she didn't have much conversation. What was happening? How could we involve her more in this?' 'Janey had most conversation about 1pm, what was happening then? Can you do that more often?'

You can also see that goals are set, and families achieve stars for improvement, regardless of whether their overall total is already very high or very low.

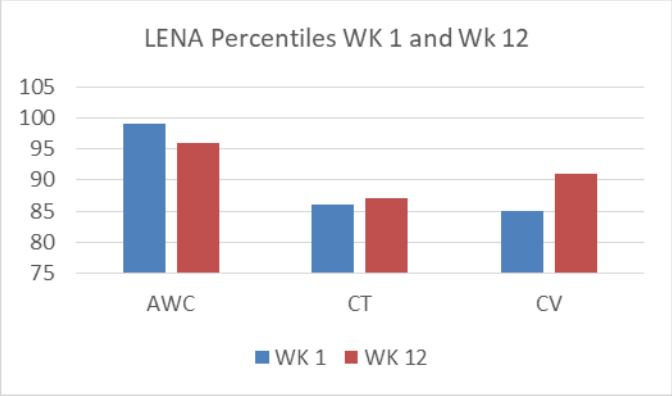
Family Feedback: 'It has been really interesting to see the difference in her and how much more responsive she has become. The confidence I have gained through this project I am now applying to other aspects of my life. I have recently become a volunteer for a radio station.' *Mum*

Coordinator Feedback: 'Mum has been so pro-active throughout the project and has also found it useful to use the journals. Mum built up a great relationship with the volunteer and found all the support so valuable. Mum has a great relationship with her daughter and since her confidence boost mum has really enjoyed singing and reading books with her daughter.' *Home-Start LENA Volunteer*

Luke

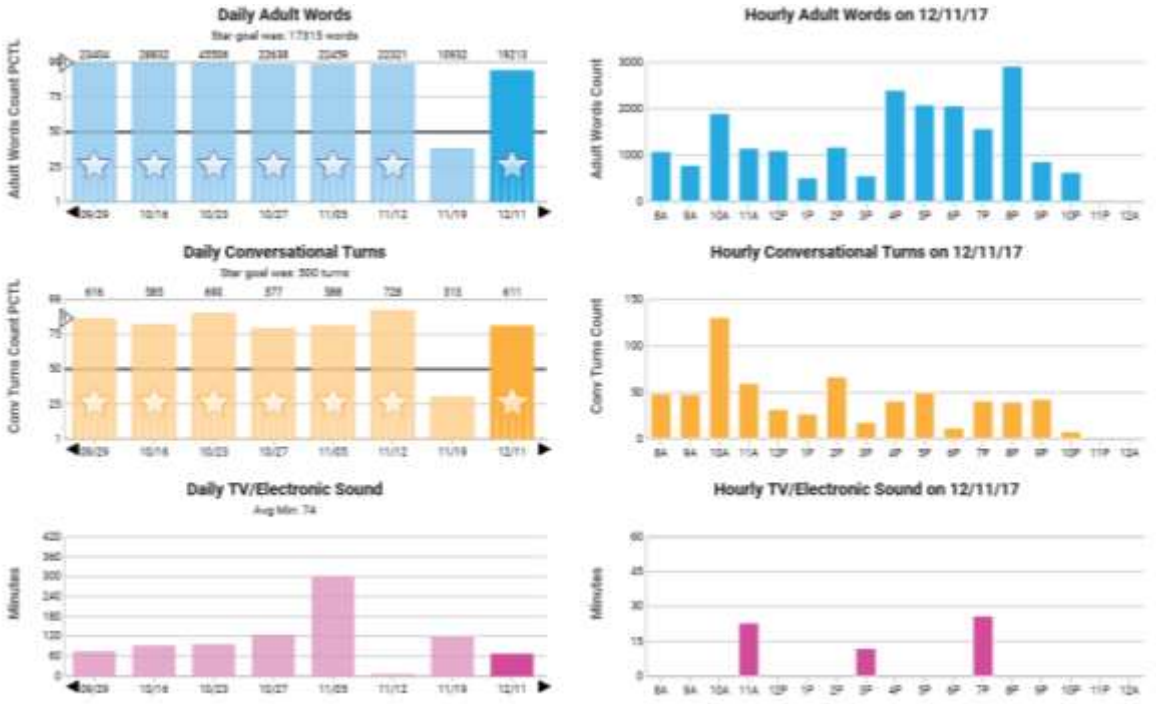
Luke is 20 months old at the start of the project. He is the third child in his family. Luke has some minor medical concerns, to which his mum is very attentive. There are toys around at home, but mum appears preoccupied with Luke's wellbeing. Mum is a qualified dental nurse but is not currently working as such. English is not the only language spoken in the home, although mum is very competent speaking with me in English. *HSUK SLT*

Outcomes:



As in the previous case study, below is the print out Luke received on week 8 of 12.

Luke's data shows consistently high numbers of words heard across the first 6 weeks of the project. In week 7 there was a significant dip. The Home-Start LENA volunteer established that this was because the family had been poorly that week. They commented at the time that they knew they had spoken less, but it was interesting to see it in the graphs the following week! Luke's family were very motivated and had good reason for fewer words on week 7 but discussing a big dip in results requires a particular sensitivity and skill on the part of the volunteer. The results are clear, and cannot and should not be hidden, but the family must still be encouraged and motivated to continue. In this case the volunteer will have picked out for praise the still low electronic sound in week 7 and gone on to talk about the high points in the hourly breakdowns in the charts on the right. As it was an unusual week, positive feedback on achieving these peaks in adversity carried them through and results were high again in week 8.



Family Feedback: 'I have really enjoyed doing the project and It has made me realise how much I do talk to my children. My little boy is very confident and happy, and I think it's because I spend a lot of time talking to him and helping him learn. LENA is a great project and I am glad I did it as it has helped me in lots of ways. I now read a lot more to my children and we try to visit the library once a month.' *Mum*

Coordinator Feedback: 'Mum had support from a volunteer throughout the project and really enjoyed the reassurance that was given by the volunteer. Mum always took on board any suggestions and feedback most weeks with the outcomes. Regardless of challenging circumstances mum always used a running commentary throughout the day and this was also witnessed on several visits.' *Home Start Volunteer*

Further quotes from families about the study may be found in Appendix C.

ANALYSIS 30: The comparison group; AWC

The US-based LENA team provided LENA normative data from a group of 21 children to our group matched on maternal education, gender and age. Unsurprisingly there were no statistically significant differences between these two groups on these variables. Data for the comparison group included AWC, CVC and CT scores at 4 time points approximately 4 weeks apart. To provide the same data for the Home Start group for analysis, AWC, CVC and CT scores were taken from week 1, week 5 and week 9 (also 4 weeks apart). So for the UK intervention groups these data were a subsample of the data reported above.

Data was tested for normality of distribution. The Kolmogorov–Smirnov test indicated the vast majority of variables were normally distributed ($p>0.05$) therefore parametric testing could be carried out. Repeated measures mixed ANOVA was conducted to compare Adult Word Count, Child Vocalisation Count and Conversational Turn Count scores across time for each group, and between our group receiving the Home Start intervention, and the US comparison group who did not receive Home-Start or indeed any other intervention that was recorded in the US database. In the analyses below, we report the differences over time and between groups for the LENA measures. We have no comparison data for the HOME inventory or the language measures.

Adult Word Count (AWC)

The table below displays the mean count scores for each group on AWC and the standard deviations of these scores. Effect sizes were also calculated which estimates the magnitude of effect (the substantive, real-world significance of the outcome). An effect size around 0.2 is considered small, 0.5 medium, and 0.7 large (Cohen, 1988).

Table 23: Mean AWC scores for the Home Start group and comparison group at the three time points.

Variable and Group	N	Mean	Standard deviation	Effect Size
Home Start AWC T1	21	12965.61	4239.42	0.33
Comparison AWC T1	21	11399.67	5174.12	
Home Start AWC T2	21	11813.45	4208.86	-0.43
Comparison AWC T2	21	14104.33	6086.43	
Home Start AWC T3	21	16409.67	8902.70	0.39
Comparison AWC T3	21	13431.64	5822.98	

So at time 1 the UK group has a higher average score. The pattern is reversed at time 2 and then flipped back at time 3. Note that the standard deviations are relatively high in both groups reflecting he variability to which we referred above.

Within group results

Mauchleys test of sphericity was significant ($p=0.039$) therefore we cannot assume equality of variance in within-subject data. The ANOVA found that AWC over time was non-significant

(Greenhouse Geisser F (1.73,69.33) =3.129, p=0.57), therefore there was no significant increase over time in AWC in both groups.

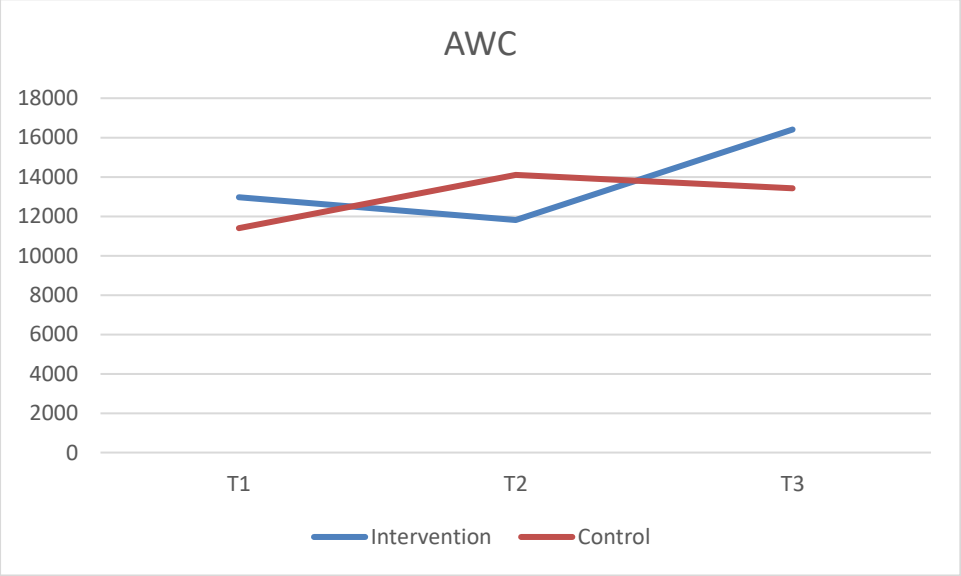
Between group results

ANOVA results indicated there was no significant Group x AWC interaction (Greenhouse Geisser (1.73, 69.33) = 2.357, p=.109). Levenes test for equality of variance was non-significant for T1 and T3, and significant for T2 (p=.039) therefore overall, we can assume equal variance between groups. ANOVA revealed there was no effect of group on AWC (F(1,40) = .908, p=.346), therefore there were no statistically significant differences on AWC data between the Home-Start and US Control group.

The effect size for AWC at T1 and T3 are small (0.3) but positive indicating great AWC in the Home Start group. However, at T2 effect size is medium (0.4) and negative, indicating greater AWC in the control group.

The graph below displays the AWC scores for both groups over time, as we can see the Home Start group begin the programme with a higher AWC than controls, with a dip in AWC towards the middle of the programme, and a rise again in AWC towards the end of the programme, ending with a higher AWC than controls.

Figure 8: Mean AWC scores for the Home Start and US comparison group over time.



ANALYSIS 31: The comparison group; Child Vocalisation Count

The table below displays the mean CV scores for both groups over time and their standard deviations and effect size.

Table 24: Mean CV scores, standard deviations and effect sizes for Home Start and US control groups.

Variable and Group	N	Mean	Standard deviation	Effect Size
Home Start CV T1	21	2034.15	739.15	0.34
Comparison CV T1	21	1728.14	1007.69	
Home Start CV T2	21	2117.25	979.69	0.47
Comparison CV T2	21	1672.86	885.81	
Home Start CV T3	21	1926.50	931.46	0.40
Comparison CV T3	21	1579.05	779.41	

Note here that the UK children had higher vocalisation rates than the US peers and note too that the standard deviations were again high.

Within group results

Mauchleys test of sphericity was non-significant (p=0.878) therefore we can assume equality of variance in within-subject data. The ANOVA found that CV over time was non-significant (F (2,78) =0.802, p=0.452), therefore there was no significant increase over time in CV in both groups.

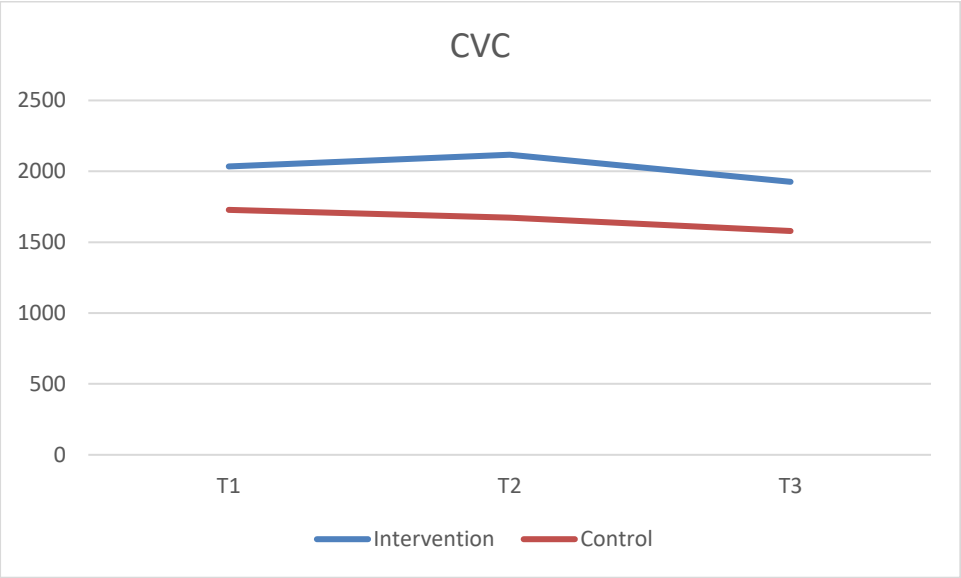
Between group results

ANOVA results indicated there was no significant Group x CV interaction (F (2, 78) = 0.164, p=.849). Levenes test for equality of variance was non-significant for T1 T2 and T3, therefore we can assume equal variance between groups. ANOVA revealed there was no effect of group on CV (F (1,39) = .2.335, p=.135), therefore there were no statistically significant differences on CV data between the Home Start and US Control group.

Effect size CV are small to medium (0.3-0.4) at all time points and indicate higher CV in the Home Start group compared to controls.

The graph below displays the CVC scores for both groups over time, as we can see the Home Start group have higher CVC than controls throughout the intervention period, with both groups slightly declining in CVC towards the end of the intervention period.

Figure 9: Mean CVC scores for the Home Start and US comparison over time.



ANALYSIS 32: The comparison group; Conversational Turn Count (CT)

The table below displays the mean CT scores for both groups over time and their standard deviations and effect size.

Table 25: Mean CT scores, standard deviations and effect sizes for Home Start and US control groups.

Variable and Group	N	Mean	Standard deviation	Effect Size
Home Start CT T1	21	439.33	175.00	0.38
Comparison CT T1	21	382.38	118.87	
Home Start CT T2	21	486.62	275.36	0.40
Comparison CT T2	21	391.19	183.29	
Home Start CT T3	21	491.00	273.13	0.40
Comparison CT T3	21	396.19	185.35	

Again we see that the UK children's turn taking was higher than that for the US groups and rises more steeply. Again the variability is quite high.

Within group results

Mauchleys test of sphericity was non-significant (p=0.376) therefore we can assume equality of variance in within-subject data. The ANOVA found that CT over time was non-significant (F (2,80) =0.10, p=0.991), therefore there was no significant increase over time in CT in both groups.

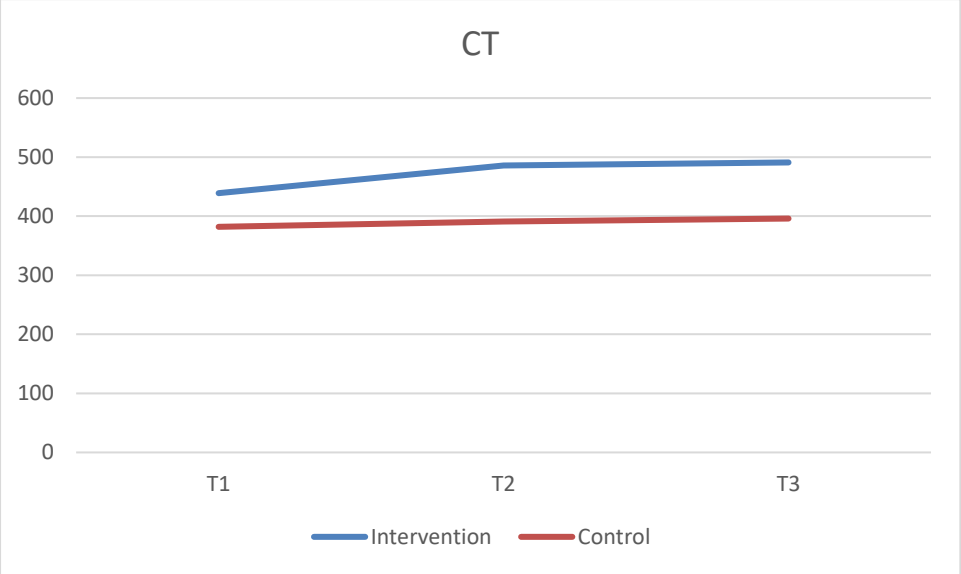
Between group results

ANOVA results indicated there was no significant Group x CT interaction (F (2, 80) = 1.300, p=.278). Levenes test for equality of variance was non-significant for T1 T2 and T3, therefore we can assume equal variance between groups. ANOVA revealed there was no effect of group on CT (F (1,40) = 1.187, p=.282), therefore there were no statistically significant differences on CT data between the Home Start and US Control group.

The effect size for CT are small to medium (0.3-0.4) at all time points and indicate higher CT count in the Home Start group compared to controls.

The graph below displays the CT scores for both groups over time, as we can see the Home Start group have a higher CT count than controls throughout the intervention period, with a slightly greater increase than controls over time.

Figure 10: Mean CT scores for the Home Start and US comparison group over time.



ANALYSIS 33: Correlations between variables

We were interested in whether the LENA variables are associated with one another within the two groups, the argument being for example that if parents are speaking more is this associated with more vocalisation on the part of the child or more recorded turn taking. One would hypothesise that these associations would be stronger in the intervention group where parents were being given messages about the associations and the need to increase communication more generally. Each LENA variable was correlated for both groups at all three time points and overall. The table below displays the correlation coefficient for each correlation (r) and significance value (p).

Table 26: Correlations between LENA variables for both groups at all time points and overall. Coefficients in bold denote significant correlations.

AWC and CT		Overall	T1	T2	T3
Home Start	r	**0.711**	**0.567**	**0.557**	**0.557**
	p	**0.000**	**0.007**	**0.009**	**0.009**
Control	r	0.116	**0.528**	0.404	**0.789**
	P	0.616	**0.014**	0.069	**0.000**
AWC and CVC					
Home Start	r	0.215	0.114	0.025	0.421
	p	0.363	0.631	0.918	0.065
Control	r	0.122	0.118	0.055	0.340
	p	0.598	0.612	0.811	0.883
CVC and CT					
Home Start	r	**0.801**	**0.786**	**0.764**	**0.830**
	p	**0.00**	**0.000**	**0.000**	**0.000**
Control	r	0.089	0.133	**0.804**	**0.748**
	p	0.703	0.564	**0.000**	**0.000**

In the Home Start group there were significant correlations overall and across all time points between AWC and CT, and between CVC and CT. In comparison, the US control group had only two significant correlations in each of these. Data suggests therefore that adult word count alone does not increase the number of child vocalisations, instead, adult words must be in the form of conversational turns to increase child vocalisations.

ANALYSIS 34: Regressions predicting LENA outcomes

Correlation and regression analysis indicated no significant correlations or regression models between the LENA AWC, CVC or CT data and the predictor variables described above – namely Child Age (months), Gender (Male/Female), Birth Order (1,2,3,4,5 etc), Maternal Education and whether the child was bilingual (Yes/No).

At this point we summarise in Table 27 which of the above within-group and between-group analyses have provided statistically significant positive results. This is following the same structure as we provided in Table 3 above.

.

Table 27 – Positive outcomes (P), Statistically significant outcomes (S)

Analyses and Variables	Within group analyses	Between group analyses	Analysis #
HOME			
Home Data; Pre- and post-intervention associations	S		1
Home Subscale Change:			
Responsivity	S		2
Acceptance	S		3
Organisation	S		4
Learning Materials	S		5
Involvement	P		6
Variety	S		7
Total	S		8
HOME Regression	S		9
LANGUAGE			
Language Test Data; Pre- and post-intervention scores	S		10
Language Test Data; Difference between time points:			
Comprehension raw score	S		11
Comprehension standard score	P		12
Comprehension percentile score	P		13
Expression raw score	S		14
Expression standard score	P		15
Expression percentile score	P		16
Language Total raw score	P		17
Language Total standard score	P		18
Language Total percentile score	P		19
Language Test Regression	P		20
LENA			
Developmental Snapshot (Age)	S		21
Boxplots	P		22
Descriptive Statistics	P		23
Change Over Time; Adult Word Count (AWC)	S	P	24
Change Over Time; Conversational Turns (CT)	P	P	25
Change Over Time; Child Vocalisation Count (CV)	P	P	26
Change Over Time; Automatic Vocalisation Assessment (AVA)	P		27
Change Over Time; TV (Secs)	P		28
LENA Case Studies	P		29
The Comparison Group			
Adult Word Count (AWC)		P	30
Child Vocalisation Count (CV)		P	31
Conversational Turns (CT)		P	32
Correlations Between Variables	S	P	33
LENA Regression	P		34

DISCUSSION AND CONCLUSIONS

This report provides data for the LENA Home-Start pilot project. This was a relatively small sample of children identified in four sites in England through local Home-Starts. The children were provided with weekly intervention to support communication within the family. The children's HOME and language scores were repeated before and after the intervention, but a special feature of the project was that the children wore a speech monitoring device developed precisely for this type of project by the LENA Foundation. The data from the four English sites was then supplemented with control group data taken from the US LENA standardisation sample; children who received no intervention for the same time period.

The overall result is, as might be expected, mixed. There are a number of examples where key variables change over time at statistical levels of significance but there are also a number that seem to show no indication of change or where the variability across children is so great that detecting a "signal" attributable to the intervention proved impossible. So, home environment and parental input increases while child measures do not clearly do so. The LENA Developmental Snapshot appears to be picking up on change for age and raw score. The illustrations of what is happening week on week from the LENA data indicate where the problem lies for some of these measures. There is a high degree of variability week on week and this creates "noise" in any analyses and reduces the chance of obtaining a statistically significant outcome.

The Language and HOME assessments

The before and after assessments were appropriate for the needs of the project. It is important to observe that the HOME Measures seemed to be especially sensitive to the type of intervention provided during the course of the intervention. Clearly the environment is mutable. One interesting feature of the children's performance on the Preschool Language Scale is that their raw scores increased significantly over the course of the programme, but their standard scores and percentiles did not, and neither did their total test scores. While this is positive of course one could reasonably argue that these children are just a little bit older and this is reflected in these raw-score change scores. The fact that their standard scores and percentile scores are not significantly changing suggest that they are not changing relative to other children in the original standardisation sample. It is important to note that the standard deviation of the two significant results was much lower than those for the other scores.

Unsurprisingly, regression analysis indicated age of child to be the strongest predictor of the language assessment (both language comprehension and expression). Interestingly in the HOME assessment, maternal education was found to be the strongest predictor of the how many learning materials were in the home, and the strongest predictor of HOME Total score was maternal education. Therefore, maternal education has a strong impact on a child's language environment; the more educated a mother is, the richer the stimulation and support for language and learning is in a child's home environment.

The LENA assessments

The Developmental Snapshot scores significantly increased over time indicating children's language ability significantly improved over the course of the programme. Within the Home Start group, parents/guardians significantly increased their word count throughout the course of the project. This is an important finding as it reflects what the volunteers were essentially asking parents to do; to talk to their child more. However, the variability or 'noise' in LENA data, as count data changed from week to week, meant that Conversational Turn counts and Child Vocalisation counts did not show significant change or increase over time. Interestingly however, within the Home-Start group positive significant correlations were found between Conversational Turns and Child Vocalisations, and between Adult Word Count and Conversational Turns, but not between Adult Word Count and Child

Vocalisations. This therefore suggests that an adult simply talking to their child more has little impact on increasing child vocalisations (i,e, more talking does not equal more turn taking). Rather the data suggest that adult talk must be in the form of conversational turns to increase child vocalisations. In the comparison US control group, the same patterns of correlation are evident between these variables, supporting this hypothesis ie that the form of adult talk is important, yet the US correlations are not evident across the board of data time points and appear a little more sporadic than in our UK data. When we compared the Home-Start group to the US control group on three time points four weeks apart, we found no significant group differences on LENA variables. This may be due to unidentified confounding variables in the US cohort, or the fact that US recording data were cited as 'approximately' four weeks apart (in fact they may have recorded closer together or further apart). In addition, the variance in our UK dataset between weeks meant heterogeneity in data was high, making it more difficult to detect significant group differences.

The challenge of measuring interaction

Measuring the detail of any interaction between speakers is a challenge. From the identification of " "motherese" ,"parentese" or what is now called "child directed speech" researchers have developed different methods of recording and transcribing the way we interact with our children. Three issues run through these developments. The first is the technology, the second the transcription and analysis of the data and third is the extent to which the context is controlled. Digitalisation has solved the first. The second is almost invariably very time consuming and thus too expensive for those engaged with service delivery. To some extent LENA overcomes this because the electronic "capture" of what is going on and rendering of that interaction into specific data takes much of the work out of the process. The risk here is the reliability of the electronic coding. Is the machine measuring what it says it is measuring? An example from another project is that the LENA recorded a conversation held in a completely tiled kitchen as TV because the tiling presumably reduced the frequency range and LENA read it as the sound of a TV. The context is a more subtle matter. Traditionally recordings were made in controlled environment with a fixed set of toys and parents are asked to "play with your child as you would do normally". Recordings are then made for thirty minutes and the first and the last ten minutes dropped from the recording that was analysed. The problems with this are in many ways obvious. To what extent is the child (and their parent) affected by being in a clinic or lab and how inhibited does the parent feel about doing what they normally do, if indeed they normally play with their child at all. The advantage with LENA, of course, is that it is effectively a free field recording and there is no control at all. So it is as naturalistic as it can be (assuming that the child forgets they are wearing the body worn recorder). The problem is that comparing individual children or indeed groups of children (as we have done above) assumes that there is some generalisability across different days – do parent and child communicate as much on the Wednesday as the Thursday. If the Wednesday is very stimulating because they went to the zoo and Thursday far less so because the parent has to go to the doctors and the child spends half the day in a waiting room playing a computer game, then LENA will pick up the variability but will have no sense of what might have been the cause of the difference. If we look at the graphs of the individual children there is a tremendous amount of normal fluctuation from week to week and this pattern differs between individual parent child dyads. This may be because the parent is or is not doing what they were asked to do by the volunteer but realistically there are so many thing going on in children lives that might influence parent child interaction and which may be much more important than the advice of a visitor. So there is a lot of noise in the signal from LENA which can make interpretation challenging. So in some was LENA is a victim of its own success, the operational challenges of measuring interaction have given way to challenges in their interpretation unless we have a complete rerecord of what was happening on the day of the recording and balancing this across the group. This, of course, may "wash out" in studies with large numbers of children and could be overcome to some extent if everyone was asked to spend a day doing similar activities or if LENA was accompanied by a home diary during the period of the intervention to shine a light on what was happening in the home from day to day.

Methodological limitations and possible solutions

Although the LENA sample was matched to the UK sample by an independent member of the LENA team and is thus effectively blind to the children involved in the intervention sample this is not true for the Home and the Preschool Language Scale which were carried out by a member of the research team in the knowledge that all children received the intervention. This may impact on post-intervention test scores. In future the second assessment should be carried out by someone who does not know the status of the children to eliminate this bias.

One of the key issues when monitoring change is the performance of the control group. It is not enough, especially with young children, to say that they seem to be improving against a background of normal developmental change. Rather we need to show that children get better relative to children of a similar level who did not receive the intervention. We attempted to do this using data from the US LENA database and received a set of data for matched children across a similar time frame. Notably, this 'similar' time frame was referred to as recordings that were 'approximately' four weeks apart from each other. Therefore, we do not know precisely how matched these recordings were to our UK data in terms of recording date/time frame. While the group was well matched for age, gender and maternal education, two of the control group could not be matched on birth order. The fact is that it would be better to have children randomly allocated to these very individualised interventions and that change can be compared over genuinely comparable groups in the same context.

Finally, we need to comment on the role of the volunteers. This is a novel element characteristic of the Home-Start model of service delivery and clearly has tremendous potential as an economical model of service delivery. One could also argue that it has potential as an "asset based" model of provision which builds on the strengths of the local community to support its more vulnerable members. The challenge is that it can make it difficult to control quality and adherence beyond the initial training. As we have seen in the introduction, the range of experience of the volunteers spanned a nanny and a doula on the one hand with presumably limited experience of language intervention to those with existing specialist skills such as a speech and language therapist on the other. This need not be a problem, but it is critical that we know more about volunteer adherence to the programme to tell whether the variability is affecting he outcome. Parent child interaction programmes which have shown consistent positive effects such as enhanced milieu therapy (Roberts and Kaiser 2011) obtain in consistent statistically significant results with moderate effect sizes of 0.4 with children with marked interaction difficulties but they do so by providing a well-articulated and consistent model of intervention which is transparent and verifiable and delivered in relatively controlled conditions. The very 'community' nature of the present study carried out in the home with the parent taking responsibility for the intervention between sessions is both its strength and its potential weakness.

Directions for future research

Clearly the lack of randomisation represents a challenge to the interpretation. One might argue that this would be unethical – ie not giving the control group children intervention but the data in the present report make this sort of argument difficult to sustain because there clearly are some children who do not appear to benefit from the intervention hence bringing down the average and thus the chances of statistical significance. In such circumstances it would be possible to use what is sometimes known as a cross over design where children not receiving in the first wave do so in a second wave while those who receive intervention it the first wave do not do so in the second. Or in a population study the stepped wedge methodology has been used to rolled out an intervention using the steps as phases where the intervention is rolled out to different groups of families and children.

In addition, a future study should provide a clearer picture of what actually took place during the intervention phase. The most appropriate method here would probably be an accompanying diary that the parent completed perhaps electronically. There are then the questions of how the parents perceived both LENA and the training. It would also be helpful to know more about what the

volunteers thought about what they were doing. And finally it would be helpful to ask parents what they understood by what they were doing both before and after the intervention. So “before and after”questionnaires to both groups would be helpful as would a diary element allowing parents to record what they were doing and what factors influenced decisions they made on a day to day basis.

Conclusions

The assessment data reported here suggest that it may be easier to detect changes on the Home environment and the adult word counts than it is in in the child specific behaviours. This probably reflects the LENA-Home programme emphasis on parental behaviour. The LENA data provide an incomparable level of detail about the parent/child interaction process at regular intervals but this detail, in this study at least, demonstrates fairly clearly that there is a great deal going on in these patterns of interaction of which the LENA Home intervention is only one element. There is potential to develop this further, but we would suggest that we need to better understand the parental perspective on the intervention and about the contribution of the volunteers. It would be important to have more feedback from the volunteers about exactly what the messages were that they were giving in the home and how they felt the parents and carers responded to the feedback and advice. The biofeedback element of LENA provides both an unrivalled level of detailed data and a powerful tool for instruction at an individual level, but these results would suggest that we are some way off translating this individualised model of intervention into results showing that the intervention works at a group level.

REFERENCES

WANG, Y. HARTMAN, M. AZIZ, N. A. A. ARORA, S. SHI, L. & TUNISON, E. "A Systematic Review of the Use of LENA Technology." *American Annals of the Deaf*, vol. 162 no. 3, 2017, pp. 295-311. *Project MUSE,*

ZIMMERMAN,I.,E., POND, R.E. & STEINER, V.G. (2014) *Preschool Language Scale - Fifth Edition (PLS-5 UK)* London: Pearson Education

GANEK, H.& ERIKS-BROPHY, A. (2018) Language ENvironment analysis (LENA) system investigation of day long recordings in children: A literature review Journal of communication disorders 72, 77-85 https://dx.doi.org/10.1016/j.jcomdis.2017.12.005

GILKERSON, J.RICHARDS, J A. & TOPPING, K. (2017) Evaluation of a LENA-Based Online Intervention for Parents of Young Children Journal of Early Intervention 39 281-298 http://dx.doi.org/10.1177/1053815117718490

HART B. and RISLEY, T., 1995, *Meaningful Differences in the Everyday Experiences of young American Children.* Baltimore: Paul Brookes.

HOFF, E. 2012 Causes and consequences of SES-related Differences in parent-to-child Speech in M.Bornstein and Robert Bradley Eds Socio-economic status parenting and child development. (pp. 147-160). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

HUTTENLOCHER, J., WATERFALL, H., VASILYEVA, M., VEVEA, J., HEDGES, L.V., 2010, Sources of variability in children's language growth. *Cognitive Psychology*, 61, 343-365.

KLATTE, I. S. & ROULSTONE, S.(2016) The practical side of working with parent-child interaction therapy with preschool children with language impairments Child Language Teaching and Therapy 32 345-359

LEVICKIS, P., REILLY, S., GIROLAMETTO, L., ULOUMUNNE, O., WAKE, M. (2014) Maternal Behaviors Promoting Language Acquisition in Slow-to-Talk Toddlers: Prospective Community-based Study Journal of Developmental & Behavioral Pediatrics: 35, 274–281 doi: 10.1097/DBP.0000000000000056

McGILLION,M., PINE,J.M., HERBERT,J.S.& MATTHEWS, D. (2017), A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds Journal of Child Psychology and Psychiatry 58:10 pp 1122–113

ROBERTS, M.Y. & KAISER, A.P. (2011) The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis American Journal of Speech-Language Pathology,20, 180–19 DOI: 10.1044/1058-0360(2011/10-0055

ROY, P., and CHIAT, S., 2013, Language and socioeconomic disadvantage: teasing apart from delay and deprivation from disorder: In Marshall, C. (ed), Current issues in developmental disorders, Hove, Psychology Press, pp 125-150.

SCHWAB, J. F. & Lew-Williams, C. Language learning, socioeconomic status, and child-directed speech Cogn Sci 2016, 7:264-275. doi: 10.1002/wcs.

SPENCER, S., CLEGG, J. and STACKHOUSE, J., 2012, Adolescents from differing socioeconomic backgrounds: a comparison of language skills. *International Journal of Language and Communication Disorders*, 47, 274-284.

SUSKIND, D.A., LEFFEL,K.R.GRAF,E., HERNANDEZ, M.W., GUNDERSON, E.A., SAPOLICH, S.G. SUSKIND. E., LINDSEY LEININGER,L., SUSAN GOLDIN-MEADOW , S. & LEVINE,S.C. (2016) A parent-directed language intervention for children of low socioeconomic status: a randomized controlled pilot study J. Child Lang. 43, 16 366-406**.**

TOSH, R.A. & ,SCARINCI, N. (2017)Parent-implemented home therapy programmes for speech and language: A systematic review International Journal of Language & Communication Disorders, 52 253-269.

WAKE, M., TOBIN, S., GIROLAMETTO, L., UKOUMUNNE, O.C., GOLD, L., LEVICKIS, P., ...& REILLY, S. (2011). Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's learn language cluster randomised controlled trial. BMJ, 343, d4741.

APPENDICES

A. Introduction to the LENA project within Home Start for parents
B. Home Start participant LENA graphs; outcomes over time
C. Quotes from families
D. Home Start UK LENA Training Guidelines

Appendix A: A. Introduction to the LENA project within Home Start for parents

Who will visit me?

My name is Emma Kirk. I am a qualified speech and language therapist. My job is to help with the LENA project.

I won't be able to give other advice or do any 'therapy' while I am visiting you. If you would like some speech and language therapy advice, ask your volunteer and they will talk to their coordinator. If the assessments show your child might benefit from more speech and language help, Home-Start can support you to find out what is available in your local area.

Who will contact me?

Emma will contact you to make an appointment.

Please try to keep the appointment, but if something changes don't worry, we can make another one. **It's really important that you call Emma or your local Home-Start if something changes - if you are not going to be at home, or it's not a good time for Emma to come.**

What happens to data about me and my child?

Your information will always be treated confidentially and will be anonymised. You can withdraw from the evaluation at any point and this will not have any impact on your Home-Start support.

Contact Emma:

07717 836688 ekirk@home-start.org.uk

Registered office: Home-Start UK
The Crescent, King St, Leicester, LE1 6RX | T 0116 464 5490 | E info@home-start.org.uk

Home Start
Support and friendship for families

Measuring the difference the LENA Project makes.

How you can help us

www.home-start.org.uk
Freephone: 0800 068 63 68

Home-Start's LENA Project

What difference does the LENA project make to children and families?

LENA helps parents improve talk and conversations with their children. The results are stronger families and children more prepared to succeed in school. Home-Start would like your help to show how this works so that other families can benefit too in the future.

How you can help us measure the difference for all children and families

As well as asking you to work with the LENA recording device and your Home-Start visitor, Home-Start would like to get some extra information from you and your child. This will help us to help other families in the future.

To get this information a different Home-Start person called Emma will come and visit you at home.

Home-Start is working with Newcastle University to analyse the information from the LENA project and write a report.

This will help us to tell other people how LENA works.

What is involved?

Emma will visit your house for a maximum of an hour and a half. She will come twice, once before you start using the LENA recording device and once after you have finished.

Emma will bring some pictures and toys to use with your child to see how they talk and what they understand. She will do the same things in both these visits. Even if your child is a baby Emma can still do this, it just won't take very long!

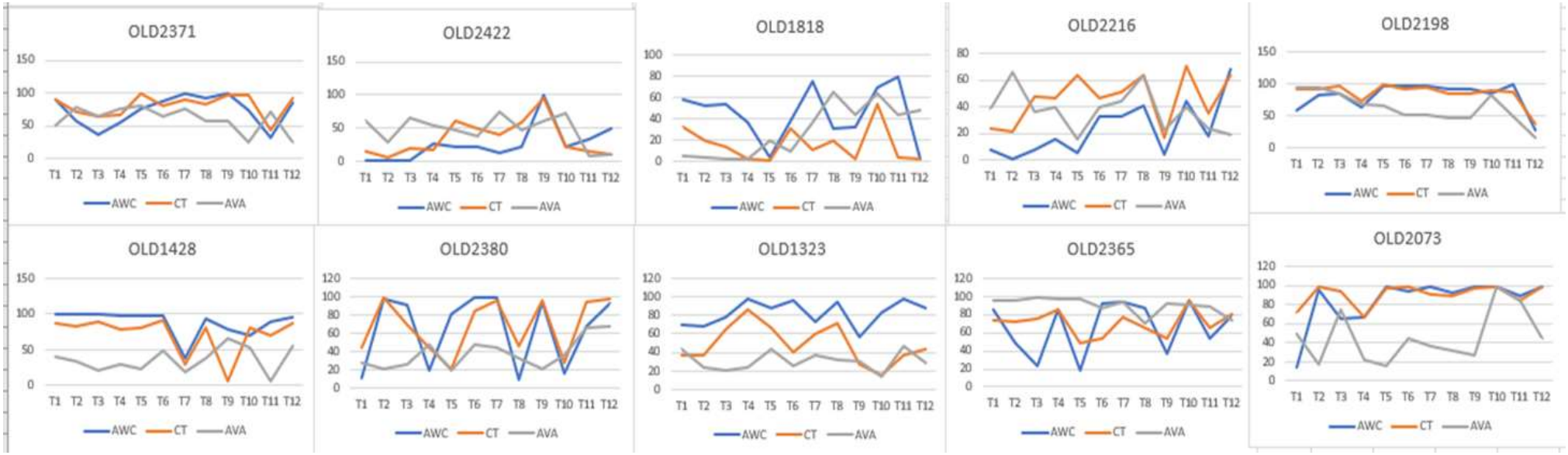
There will also be some questions for you, because you know your child best.

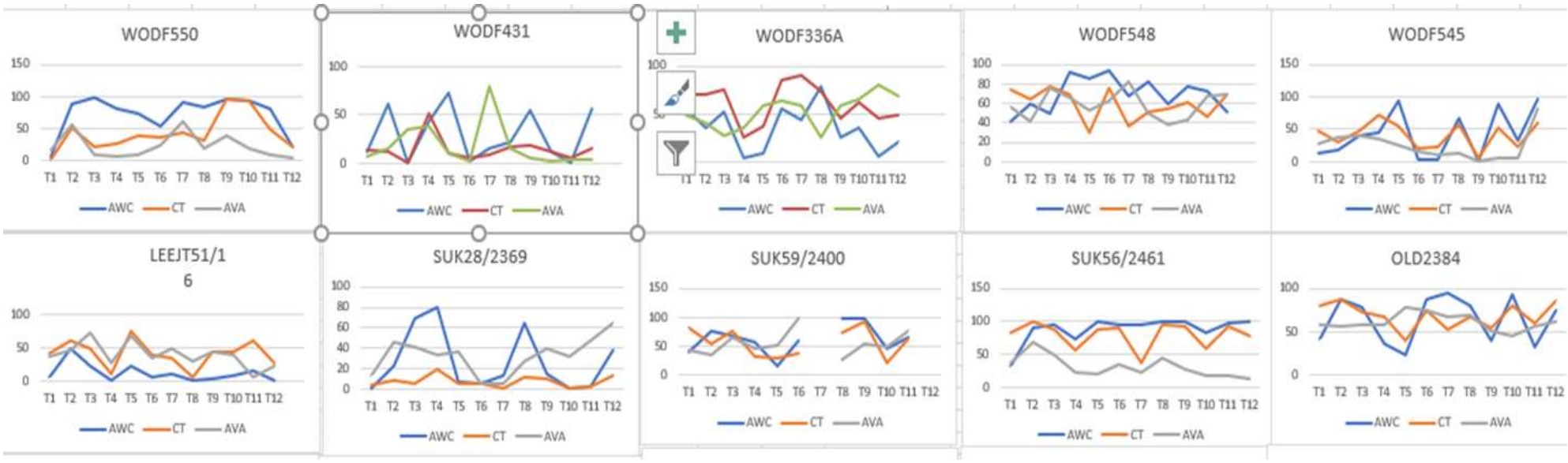
Why do the visits need to be done by another Home-Start person?

Emma will do the assessments because they will take a bit longer than a regular visit, and because she is a qualified speech and language therapist. A speech and language therapist has been chosen because she will help Home-Start to understand the things we learn from the project.

This doesn't mean your child needs a speech and language therapist.

Appendix B: Home Start participant LENA graphs; outcomes over time




WODF550
WODF431
WODF336A
WODF548
WODF545
LEEJT51/16
SUK28/2369
SUK59/2400
SUK56/2461
OLD2384
T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12
AWC CT AVA

Appendix C: Quotes from families

"LENA has helped me to identify how much time I need to play with my son"

Mum 1428

"My son seems a lot calmer and I think that's because I am taking more time to play with him"

Mum 2073

"I enjoy doing activities with my children as I know how much interacting benefits them"

Mum 1972

"I feel like a better mum because of LENA"

Mum 2384

"I didn't realise how little I actually spoke to my child, this has really helped me understand how to support her development

Mum 2455

"I feel I really understand him better"

Mum 2380

"I am really looking forward to using what I learnt with my son, with my baby daughter"

Mum 2380

"I feel like my relationship is more meaningful with my son"

Dad 2380

"It has been interesting to see how more responsive my daughter has become. The confidence I have gained through doing LENA I am now applying to other areas of my life and I am now volunteering on a radio station"

Mum 2198

"Through doing LENA I have become more encouraging with my son and because of this I think it has given him the confidence to learn to walk"

Mum 2587

"You have reassured my wife she is doing a good job"

Dad 2307

"I can really see a difference with how much more my daughter talks compared to when my other children were this age"

Mum 1323

"It's nice that my neighbour has done this as well because she is able to support me on the LENA project"

Mum 2604

"My son is partially sighted but it certainly hasn't held his speech back, LENA is proving that"

Mum 2482

HOME-START UK LENA TRAINING

The total training time is approximately 7.5 hours.
It contains 4 sessions.

The timings are based on 12 volunteers participating in the training; timings could vary with more or less volunteers.

Overall aim of the training:
To give existing Home-Start volunteers the confidence, knowledge and skills to deliver the LENA Home Coaching Programme to families.

Outcomes: by the end of the training participants will be able to:

- Describe and discuss all aspects of the LENA programme
- Explain the importance of early language development
- Deliver the LENA Home coaching programme to families

Session 1: Introduction to LENA

Aim: To introduce volunteers to the principles and practice of the LENA home coaching programme

Outcomes: by the end of the session participants will be able to:

- Explain the different elements of the LENA programme
- Describe how interactive talk grows babies brains
- Discuss the importance of interactive talk in terms of language development

Session 2: LENA in action

Aim: To give volunteers a deeper understanding of the LENA programme and how it will work with Home-Start families

Outcomes: by the end of the session participants will be able to:

- Compare and contrast the LENA programme with other forms of Home-Start support
- Explain the differences between coaching and befriending
- Use the LENA recorder and vest and be able to explain its use to a family
- Explain the content of the LENA report
- Describe the Talking Tips and develop their practical application with families
- Discuss the purpose of the LENA snapshot and how it used with families
- Discuss the 14 LENA Home Coaching Sessions

Session 3: Communicating with parents/carers on the LENA programme

Aim: To give volunteers an opportunity to explore and examine coaching, effective questioning and goal setting for supporting parents on the LENA programme

Outcomes: by the end of the session participants will be able to:

- Explain the meaning of coaching in relation to the LENA programme
- Describe the style of coaching
- Identify and use a range of coaching skills
- Describe empathy and use an empathetic approach
- Discuss the importance of helpful questioning
- Set goals with parents on the LENA programme

Session 4: Preparing to support a family on the LENA Home Coaching Programme

Aim: To ensure volunteers are fully prepared and confident to deliver the LENA Home Coaching Programme to a family

Outcomes: by the end of the session participants will be able to:

- Describe in detail the content of the first three home visiting sessions
- Plan the first three home visiting sessions
- Facilitate parent to analyse the LENA report
- Complete the volunteer weekly diary sheet